REPORT RESUMES

ED 003 059 24
STUDIES IN GUIDED LEARNING.
BY- COOK, JOHN O. MILLER, HOWARD G.
NORTH CAROLINA UNIV., RALEIGH, N.C. STATE UNIV.
REPORT NUMBER CRP-1242 PUB DATE 63
EDRS PRICE MF-$0.18 HC-$4.84 121P.

DESCRIPTORS- *LEARNING PROCESSES, *COLLEGE STUDENTS, *SEQUENTIAL LEARNING, SERIAL ORDERING, TRANSFER OF TRAINING, ANXIETY, REINFORCEMENT, *EDUCATIONAL EXPERIMENTS, SEQUENTIAL LEARNING, *PROMPTING, RALEIGH, NORTH CAROLINA

SIX SEPARATE STUDIES, ALL CONCERNED WITH GUIDED TRIALS DURING THE LEARNING PROCESS, WERE REPORTED. SPECIFIC ASPECTS COVERED BY THE RESEARCH INCLUDE--(1) GUIDANCE AND SYMBOLIC LEARNING AND (2) GUIDANCE AND SEQUENTIAL LEARNING. VARYING NUMBERS OF COLLEGE UNDERGRADUATES WERE USED AS SUBJECTS IN THE SIX EXPERIMENTAL PROCEDURES. THESE INCLUDED--(1) PROMPTING AND CONFIRMATION IN PAIRED-ASSOCIATE LEARNING, (2) THE EFFECT OF TEST SCHEDULES UPON LEARNING UNDER CONDITIONS OF GUIDANCE, (3) THE INTERACTING EFFECTS OF GUIDANCE AND ANXIETY UPON LEARNING, (4) THEORETICAL IMPLICATIONS OF RESPONSE ANALYSIS IN PAIRED-ASSOCIATE LEARNING, (5) AMOUNT OF TRANSFER AS A FUNCTION OF TYPE OF GUIDANCE IN SOLVING SERIES-COMPLETION PROBLEMS, AND (6) A GENERALIZED PLAN FOR SERIAL LEARNING. CONCLUSIONS DRAWN FROM THE RESULTS WERE APPLIED TO THE DEVELOPMENT OF A THEORY OF SERIAL LEARNING. (WB)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Studies in Guided Learning

Cooperative Research Project No. 1242

John Oliver Cook and Howard G. Miller

NORTH CAROLINA STATE of the UNIVERSITY of
NORTH CAROLINA at RALEIGH

Raleigh, North Carolina

1963

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

ED 003 059

Preface

The papers in this set of studies are all concerned in one way or another with the problem of guiding the learning process. The first paper, titled "Prompting and Confirmation in Paired Associate Learning" addresses itself to the question of what function is performed by test trials (with feedback) in a series of guided training trials. The second study, "The Effect of Test Schedules Upon Learning under Conditions of Guidance," concerns itself with the question of whether there is some optimum way to intersperse test trials (with feedback) among guided training trials. The learning task here is a paired associate one. In Experiment 3, "The Interacting Effects of Guidance and Anxiety upon Learning," a paired associate task is also used. But in this study the aim is to discover whether there is some interaction between the training method and the subject's degree of anxiety. In other words, it takes up the question of whether anxious students can learn paired associates better under one training method (guidance) than under another.

The fourth paper, "Theoretical Implications of Response Analysis in Paired Associate Learning," is a non-experimental one which argues for a finer analysis of the responses in paired associate learning as a means of testing theoretical implications. The fifth paper, "Amount of Transfer as a Function of Type of Guidance in Problem Solving," employs series-completion problems and is aimed at testing the notion that guidance in a problem solving task has a deleterious effect upon transfer to the new problem. The final paper in the set, "A Generalized 'Plan' for Serial Learning," is devoted to investigating the effects of guidance in a serial learning task. However, as the title suggests, perhaps the chief focus of interest in the paper is the development of a theory of serial learning.

TABLE OF CONTENTS

Page

PREFACE i

LIST OF TABLES ii

LIST OF FIGURES iv

EXPERIMENT A-1 1
Prompting and Confirmation in Paired Associate Learning

EXPERIMENT A-2 12
The Effect of Test Schedules upon Learning Under Conditions of Guidance

EXPERIMENT A-3 45
The Interacting Effects of Guidance and Anxiety upon Learning

STUDY A-4 64
Theoretical Implications of Response Analysis in Paired Associate Learning

EXPERIMENT A-5 70
Amount of Transfer as a Function of Type of Guidance in Solving Series-Completion Problems

EXPERIMENT B-1 79
A Generalized "Plan" for Serial Learning

SUMMARY 1-15

LIST OF TABLES

Page

EXPERIMENT A-1

Table 1. Experimental design 4
Table 2. Format for presenting items during training (prompting and test trials).. 5
Table 3. Mean number of items correct in Part I.. 6
Table 4. Number (and Percent) of perfect scores in Part I 7

EXPERIMENT A-2

Table 1. Mean number of correct responses out of a possible ten, for each of the experimental groups on test one 27
Table 2. Mean number of legitimate responses out of a possible ten, for the experimental groups elicited on test one 28
Table 3. Mean number of different legitimate responses out of a possible ten emitted on test two 28
Table 4. Results of the analysis of variance using C scores 29
Table 5. Results of the analysis of variance using D scores 30
Table 6. Means and standard deviations of the four experimental groups 35

EXPERIMENT A-5

Table 1. Median time to solution (in seconds) ...72
Table 2. Non-parametric analysis of variance for each problem 72
Table 3. Mann-Whitney U-Test comparisons between the female-GG cell and the other cells in Problem #4 ... .73
Table 4. Strategies that Ss reported using .. ...74.
Table 5. Strategy 1 (Adding Successive Numbers) in Problems #1 and #2 76
Table 6. Strategy 1 (Adding Successive Numbers) in Problem #1, and Strategy 3 (Adding Alternate Numbers)in Problem #2 77
Table 7. The effect of solving Problem #1 by Strategy 1 (Adding Successive Numbers) upon the choice of Strategy 4 (Subtracting Successive Numbers) as the initial strategy in Problem #4 77

Experiment A-5 (cont.)

Table 8. Strategy 1 (Adding Successive Numbers in Problem #1, and Strategy 4 (Subtracting Successive Numbers) in Problem #4 78

Experiment B-1

Table 1. Mean trials to criterion ... .84
Table 2. Interview data 85

LIST OF FIGURES

Page

EXPERIMENT A-2

Figure 1. An example of the paired-associate material used by Cook and Kendler (1956) 13
Figure 2. The Cook and Kendler (1956) model 14
Figure 3. Timing of intervals and duration of S_1 and S_2 terms 24
Figure 4. The specific arrangement of prompt and confirmation trials for each group... 25
Figure 5. The response sheet 26
Figure 6. The use of an obvious transposition as a cue 30
Figure 7. Sequencing of prompting and confirmation trials by the three levels of the SD group 33

EXPERIMENT A-3

Figure 1. Design for experiment. All groups contained 10 subjects each 49
Figure 2. High anxiety groups (no stress) prompting vs confirmation 51
Figure 3. Low anxiety (no stress) prompting vs confirmation 51
Figure 4. High anxiety stress prompting vs confirmation 52
Figure 5. Prompting low anxiety, high anxiety and high anxiety stress groups. The high anxiety stress group had only four testing trials 52
Figure 6. Confirmation, low anxiety, high anxiety and high anxiety stress. The high anxiety stress group had only four testing trials 53
Figure 7. Performance for all groups. Stress groups had only four test trials ... 53
Figure 8. Number of nonspecifics emitted during learning trials 55
Figure 9. Number of nonspecifics emitted during tests 55
Figure 10. Regression lines drawn on the relationship between GSR activity levels and number of correct responses for prompting and confirmation during training 56

Experiment A-3 (cont.)

Figure 11. Regression lines drawn on the relationship between GSR activity levels and number of correct responses during testing for prompting and confirmation 57

EXPERIMENT A-4

Figure 1. A multi-process model of paired associate learning ... ,.. ... ..., 65
Figure 2. An analysis of responses in paired associate learning experiments ... 66
Figure 3. An illustration of an analysis of a single S's scores in a paired associate learning experiment ... 67

EXPERIMENT B-1

Figure 1. The integration of serial learning ... 80
Figure 2. Diagram of a generalized "Plan" for serial learning 87

Prompting and Confirmation in Paired Associate Learning

John Oliver Cook, Howard G. Miller
James Brown Grier and James W. Staman

Few things in the psychology of learning are as well grounded and as widely accepted as the principle that knowledge of results facilitates learning. This principle is the rationale behind the practice of giving tests during academic courses and the practice of employing a training technique that tests the students either continually, as in a Skinner-type linear program, or at intervals.

It has also been found that a guidance procedure called prompting without knowledge of results also facilitates learning. Several experiments in paired associate learning (Cook and Kendler, 1956; Cook, 1958; Cook and Spitzer, 1960) found that prompting was superior to the procedure called confirmation. In training under the prompting procedure Ss are briefly shown the stimulus term, then briefly shown the response term and then asked to reproduce the response term. In the confirmation procedure Ss are briefly shown the stimulus term, then asked to reproduce the response term, and then are briefly shown the response term, thus giving them knowledge of results and enabling them to check their answers. Using different paired associate materials from those used in the previous prompting vs. confirmation experiments, Briggs (1961) compared a prompting procedure with three different kinds of confirmation procedure (single-try confirmation, multiple-try confirmation, and single feed-back) and found that prompting was significantly superior to all of them on a test of immediate retention. On a test given two weeks after training prompting was still superior, though in the case of one comparison (prompting vs. single-try confirmation) the significance level had dropped to .094.

Angell and Lumsdaine (1961) pointed out that the prompting condition in the experiments by Cook, et. al., was not a pure one, inasmuch as a test trial without knowledge of results was inserted after every third training trial. Their view, which seems to be shared by Skinner (1958), is that 100% stimulus support throughout training, as in pure prompting, may be undesirable. The hypothesis is that a training procedure which combines prompting trials with training trials that require S to respond to the stimulus alone (without the prompt) may produce better results than pure prompting. In their experiment, Angell and Lumsdaine compared a pure prompting treatment with a prompting-plus-confirmation treatment in which a confirmation trial was substituted for every fourth prompting trial. All Ss were given 24 training trials and were tested immediately after the twelfth and twenty-fourth training trials and again fifteen minutes later. Using the number of persons achieving perfect scores as the dependent variable, Angell and Lumsdaine found the P plus C treatment to be superior to the P treatment on all three tests, though the closest the difference came to significance was a p value of .10 on test three; that is, on the test given fifteen minutes after training. However, combining the results of all the tests yielded a p value of less than .05.

The main purpose of the present experiment was to try to find out what facilitative function, if any, is performed by confirmation trials inserted periodically in a series of prompting trials. There are several possibilities. If we give S the correct answer after he has made his response, as in a confirmation trial, the confirmation trial may be facilitative simply because it provides S with another opportunity to learn the correct response. If S is not shown the correct answer but is simply told later "right" or "wrong", the confirmation trial may be facilitative by virtue of the fact that it tells S what he does not know, thus enabling him in subsequent training trials to concentrate his efforts on learning what he does not yet know. If S is not given any feed-back at all, the test trial (we can hardly call it a confirmation trial, since nothing is confirmed or disconfirmed) may be facilitative simply because it motivates S. Another possibility, of course, is that confirmation trials are not facilitative at all.

Method

Materials.- The paired associate materials used in this experiment were the same as those used in experiments by Cook and Kendler (1956), Cook (1958), Cook and Spitzer (1960) and Angell and Lumsdaine (1961). The paired associate list consisted of 10 items. The stimulus terms were the first 10 letters of the alphabet and the response terms were "code lines", each consisting of two straight lines that connected 3 dots of a standard, non-representational, seven-dot pattern. The materials are pictured in Cook (1963). A trial consisted in presenting each of the ten items once. The items were presented in a different random order on each trial.

Experimental Design.- Although Ss were randomly assigned to all seven conditions, the experimental design can be most easily understood as consisting of two parts. Training trials of the prompting type (indicated by a "p" were used throughout both parts of the experiment. The numbers in Table 1 between adjacent trials indicate the number of seconds that elapsed between those trials. For example, after the third prompting trial there was an interval of 300 seconds. This interval was necessitated by the fact that it was necessary during this interval to give test instructions to Ss in the first three conditions of Part I of the experiment. Again, there is an interval of 120 seconds just preceding the final test trial. In all other cases, whether the trial was a training trial or a test trial, the intertrial interval was 10 seconds.

In Part I of the experiment there were 5 conditions. In the T_0 condition after every third prompting trial there was a test with no feedback. That is, the subjects were shown each stimulus term and were required to produce the response but were given no information about the correctness or incorrectness of that response. In this condition, as in the other conditions of Part I, a final test (T_F) was administered 120 seconds after Test 5. Condition T_{RW} was exactly the same as condition T_0 except that on each of the first 5 tests S was informed after making each response whether the response was right or wrong.

Table 1

Experimental Design

Part I																								
Condition	**N**																							
T_O	10	P	P	P	300	T_O^1	P	P	P	T_O^2	P	P	P	T_O^3	P	P	P	T_O^4	P	P	P	T_O^5	120	T_F
T_{RW}	10	P	P	P	300	T_{RW}^1	P	P	P	T_{RW}^2	P	P	P	T_{RW}^3	P	P	P	T_{RW}^4	P	P	P	T_{RW}^5	120	T_F
T_C	10	P	P	P	300	T_C^1	P	P	P	T_C^2	P	P	P	T_C^3	P	P	P	T_C^4	P	P	P	T_C^5	120	T_F
P_A	5	P	P	P	300	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	120	T_F
$(P+O)_A$	5	P	P	P	300	0	P	P	P	0	P	P	P	0	P	P	P	0	P	P	P	0	120	T_F
Part II																								
P_B	5	P	P	P	300	P	P	P	P	P	P	P	P	P	120	T_F								
$(P+O)_B$	5	P	P	P	300	0	P	P	P	0	P	P	P	0	120	T_F								

Condition T_C was exactly the same as the two other conditions except that on each of the first five tests, S, after making each response, was shown the correct response. The T_C condition corresponds in general to the P + C treatment in the experiment by Angell and Lumsdaine (1961). The P_A condition corresponds to the pure prompting condition in the Angell and Lumsdaine study. In the $(P+O)_A$ condition an empty time interval equal in length to a test and to a prompting trial was inserted after every prompting trial.

The two conditions in Part II of the experiment were the same as the last two conditions in Part I except that the final test was administered earlier in order to provide some hint about the shape of the learning curve under these two conditions.

The formats for presenting items during test trials are shown in Table 2. The Table is self-explanatory.

Table 2

Format for presenting items during training (prompting) and test trials

Seconds

Type of Trial	1	2	3	4	5	6	7
P	Saw Letter	Saw Code Line		Drew Code Line			Interval
T_O	Saw Letter	Drew Code Line			Interval		Interval
T_{RW}	Saw Letter	Drew Code Line			Told "Right" or "Wrong"		Interval
T_C	Saw Letter	Drew Code Line			Saw Code Line		Interval
T_F	Saw Letter	Drew Code Line					Interval

Apparatus.- The apparatus used in this experiment was similar to that employed in the previously cited experiments using the same materials. It consisted of a film strip projector which automatically projected frames onto a screen. The timing was accomplished automatically by means of 300 cps beeps recorded on ordinary audio tape. The beeps activated a switching relay which advanced the film after the appropriate interval. The tape also carried the instructions to the subjects.

Subjects and Procedure.- The *S*s were 50 undergraduate students in an introductory psychology course. They were given 1% grade credit for participation in the experiment. *S*s were run two at a time in a small windowless room. When only one *S* failed to show up, a graduate student substituted for the missing *S* unbeknown to the real *S*. In the T_{RW} condition the words "right" and "wrong" were spoken by one of the *E*s. All *S*s practiced covertly during the training trials. During the test trials *S*s responded overtly by drawing the code lines on sheets that had printed on them the standard seven-dot pattern.

Results

The mean number of correct items in Part I of the experiment is shown in Table 3.

Table 3

Mean Number of Items Correct in Part I

Condition	*N*	Test Number					
		T^1	T^2	T^3	T^4	T^5	T_F
T_O	10	3.0	4.6	6.7	8.5	9.1	8.9
T_{RW}	10	2.5	5.1	7.2	8.4	9.5	9.4
T_O	10	2.0	4.5	6.2	7.2	8.4	8.8
P_A	5	-	-	-	-	-	8.2
$(P+O)_A$	5	-	-	-	-	-	8.6

None of the differences on any of the tests is significant at the .05 level by either a parametric or nonparametric test. The means in Part II of the experiment are 8.0 for the P_B condition and 6.2 for the $(P+O)_B$ condition. This difference is significant at the .05 level.

Angell and Lumsdaine (1961) used the number of perfect scores as the dependent variable rather than means or medians. The number (and the percent of the total) of perfect scores in Part I of the present experiment are shown in Table 4.

Table 4

Number (and Per Cent) of Perfect Scores in Part I

Condition	N	T^1	T^2	T^3	T^4	T^5	T_F
T_O	10	0(0)	0(0)	3(30)	4(40)	3(30)	4(40)
T_{RW}	10	0(0)	1(10)	4(40)	5(50)	6(60)	8(80)
T_C	10	0(0)	0(0)	2(20)	2(20)	5(50)	6(60)
P_A	5						1(20)
$(P+O)_A$	5						2(40)

None of the differences on any of the tests is significant at the .05 level by a chi-square median test. In Part I of the experiment the numbers of perfect scores were 3 (60%) for the P_B condition and 0 for the $(P+O)_B$ treatment. This difference is significant at the .05 level by a chi-square median test.

It should be noted that the correlation between means and number of perfect scores, while high, is not perfect. In the present experiment the product-moment correlation over all the tests is .89. The discrepancy between these two dependent variables has an important implication. Note the scores on the final test (T_F) for the T_O condition and the T_C condition in Tables 3 and 4. If we were employing a one-tailed statistical test and using the means we would have to reject the hypothesis that T_C is greater than T_O, but if we were using the number of perfect scores we would have to reject the hypothesis that T_O is greater than T_C.

Discussion

There are at least four, and possibly five, differences between the Angell and Lumsdaine (1961) study and Part I of the present experiment: (a) The format for item presentation. Angell and Lumsdaine have a .257-second interval after the presentation of the stimulus term (letter). The present experiment has no interval except for the time that it takes for the film strip projector to change the frames. (b) Number of training trials. Ss in the Angell and Lumsdaine experiment were run for 24 trials, and in the present experiment for 20. (c) Location of tests. In the Angell and Lumsdaine study Ss were tested after the 12th trial, after the 24th trial, and again 15 minutes later. In the present experiment some groups were tested after every third prompting trial and again 2 minutes after the fifth test. Other groups were tested after the end of training, and again two minutes later. (d) Different Ss. Angell and Lumsdaine used high school students, college students and non-students, all of whom were paid volunteers. The present experiment used unpaid college students who were given grade credit for participation in the experiment. (e) Inter-trial interval. The intertrial interval in the present experiment was 10 sec. In the Angell and Lumsdaine study the inter-trial interval is not reported.

In spite of these differences in procedure the results of the present experiment confirmed the finding of Angell and Lumsdaine of no significant difference on tests given during, or immediately after, training between a pure prompting condition (P_A) and a condition (T_C) in which a confirmation trial is substituted for every fourth prompting trial. On the other hand, as far as tests given during training and immediately afterwards are concerned, the present experiment provides no support for the suggestion that "a training sequence that incorporates some unprompted trials will be more effective than one that does not" (Angell and Lumsdaine, 1961, page 397). However, future research may disclose that, as far as retention over a period of time is concerned, some combination of prompting and confirmation trials is superior to a straight prompting procedure, perhaps in somewhat the same way that partial reinforcement enhances resistance to extinction in operant conditioning.

One possible interpretation of the outcome of the present experiment is that the lack of inter-group differences in the dependent variable points to a constant factor among the independent variables, namely the total running time for each S. Bugelski (1962) and Bugelski and Rickwood (1963) have presented evidence that "Presentation time x Trials (or "Total" time) is constant for learning paired associate nonsense syllables" (Bugelski and Rickwood, 1963). The results of the present experiment seem to enable us to extend Bugelski's generalization. Bugelski, and Bugelski and Rickwood held training method constant for all Ss: they used an anticipation method. In the present experiment we have used predominantly a training procedure (prompting) which is different from the one used by Bugelski, and Bugelski and Rickwood, and, moreover, we have varied the training procedure from group to group by replacing some training trials with test and others by empty intervals. Yet we have come out with the same results as Bugelski, and Bugelski and Rickwood. However, the finding of a significant difference between the two conditions in Part II of the experiment limits the scope of the generalization. To put it rather cautiously, we can say that given enough training trials - though we cannot at present specify how many are enough - the total learning time per S seems to be an important variable in paired associate learning. This generalization holds for at least two different training methods (anticipation and prompting) and even for a certain mixture of training methods. (Condition T_C is a mixture of prompting and confirmation trials).

These remarks should not be interpreted as saying that Ss learn a given list of paired associates in a fixed amount of time regardless of training methods. It has been repeatedly found (Cook and Kendler, 1956, Cook, 1958, Cook and Spitzer, 1960, Briggs, 1961) that a prompting procedure produces better results than a confirmation procedure even though stimulus exposure times, the duration of all intervals and total training time are all simultaneously held constant. Thus, while the constant time principle holds even though there is some variation in the training method, it cannot be said to hold over all training methods; nor was it, we believe, Bugelski's intention to assert that it did hold regardless of training method.

The papers by Bugelski (1962) and Bugelski and Rickwood (1963) suggest, if they do not explicitly assert, that during paired associate training Ss practiced during the intervals. If the evidence, so to speak, forces us into making this assertion then we are faced with an interesting question, namely, what do the Ss practice during the intervals? Do they practice the gamma connection? That is, do they practice connecting the implicit response to the response term to the stimulus properties of the implicit response to the stimulus term? Since both the cue and the response, in this case, are internal, this is a possibility. Do they practice the alpha and beta processes? That is, do they practice making implicit responses to the two external stimuli? Is this possible? In other words, is it possible to practice a response to an external stimulus in the absence of that external stimulus? Providing suitable answers for these and similar questions will involve us all in a great deal more work, in the course of which we will no doubt learn more about paired associate learning than we know at present.

References

Angell, D., and Lumsdaine, A. A. Prompted and unprompted trials versus prompted trials only in paired-associate learning. In A. A. Lumsdaine (Ed.) *Student response in programmed instruction*, Nat'l. Acad. Sci.-Nat'l. Res. Coun. Publ. 943, Washington, D. C., 1961, 389-398.

Briggs, L. J. Prompting and confirmation conditions for three learning tasks employing the subject-matter trainer. In A. A. Lumsdaine (Ed.) *Student response in programmed instruction*. Nat'l. Acad. Sci.-Nat'l. Res. Coun. Publ. 943, Washington, D. C., 1961, 375-388.

Bugelski, B. R. Presentation time, total time, and mediation in paired-associate learning. *J. exp. Psychol.*, 1962, 63, 409-412.

Bugelski, B. R., and Rickwood, J. Presentation time, total time, and mediation in paired-associate learning: Self-pacing. *J. exp. Psychol.*, 1963, 65, 616.

Cook, J. O. Supplementary report: Processes underlying learning a single paired-associate item. *J. exp. Psychol.*, 1958, 56, 455.

Cook, J. O. and Kendler, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch and F. Cameron (Eds.) *Symposium on Air Force human engineering, personnel, and training research*. Nat'l. Acad. Sci.- Nat'l. Res. Coun. Publ. 455, Washington, D. C., 1956, 90-98.

Cook, J. O., and Spitzer, M. E. Supplementary report: Prompting vs. confirmation in paired associate learning. *J. exp. Psychol.*, 1960, 58, 275-276.

Skinner, B. F. Teaching machines. *Science*, 1958, 128, 969-977.

The Effect of Test Schedules upon Learning Under Conditions of Guidance

James Wright Staman and John Oliver Cook

Introduction

Psychology's pioneer in the field of automated teaching is S. L. Pressey. As early as 1924, he exhibited a "simple apparatus which gives and scores tests - and teaches" (Pressey, 1926, 373). However, it was not until Skinner's now historic article of 1954 was published, "The Science of Learning and the Art of Teaching," that educators and psychologists seriously concerned themselves with the possibilities of such devices.

Since 1954, however, there has been a growing volume of literature, some describing designs for self-instructional devices, others theorizing as to the optimum procedures for programming the material to be learned, and other reports of research designed to test the effect of specific variables. Schramm (1962), for example, notes that there have been approximately 190 reports of original research on programmed instruction since 1954, and that more than 165 of these have appeared since 1959.

Lumsdaine (1960) has delineated several major classes of questions which arise in the programming of material. These include: (1) the size of the "step" or length of frame, (2) content of correction or answer frame, (3) the logic of programmed instruction (e. g., uniform versus contingent sequences), and (4) content of prompts or manner of cueing.

Within the last category, there are several questions to which research might be addressed. One of them is: What is the general validity of the hypothesis that there should be a transition from maximum cue strength? Can this principle be extended or made more precise in terms of some quantitative parameters for optimal cue-strength gradients or supplementation in terms of auxiliary principles? (Lumsdaine, 1960, 538-539) The present study is designed to provide information which will aid in answering this question.

Review of Literature

In the course of investigating the effectiveness of audience participation on performance on subsequent tests (Kendler, Cook, Kendler, 1956; Kendler, Kendler, Cook, 1954), a study was done to determine the effects of prompting during a review session. After viewing a film which explained the meaning of twenty conventional map signs, all subjects were given a review session. Their task was to call out the name of the map sign as soon as it appeared. Prompting consisted of having the experimenter call out the name of the sign as soon as it appeared. Half the subjects were prompted, the other half were not. The performance of the prompted group was superior to that of the unprompted group. The difference, however, was not statistically significant. Of practical value was the fact that the study led to an exhaustive analysis of the effects of audience participation on subsequent test performance. In every instance where audience participation had been shown to be superior to non-participation, the subject's participation responses were confirmed by providing feedback information.

Two possible explanations of these findings were suggested.

One possibility was that the confirmation of responses acted as a reinforcement for the response (or as knowledge of results) and thus the facilitation. The other possibility was that facilitation occurred because the responses were prompted. It was possible, for example, for some subjects to withhold their responses until they had heard the majority call out the correct response, then make their response.

In order to evaluate the effectiveness of these two procedures, a study was made which compared a prompting and a confirmation mode of practice at two levels of training: early in learning and late in learning (Cook and Kendler, 1956).

The task used was paired-associate learning of a stimulus term familiar to the S (a letter of the alphabet) and novel response (a bent code line that connected three dots in a matrix of seven randomly placed dots). (See Figure 1.)

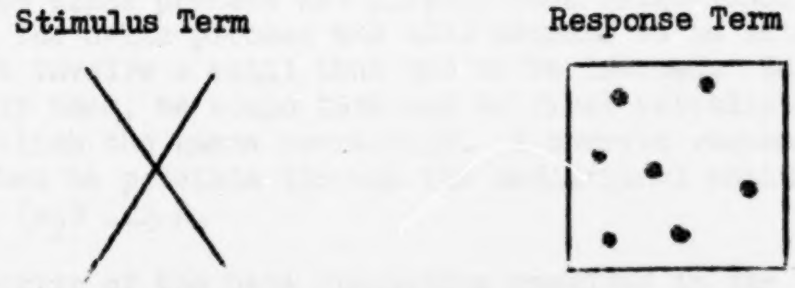


Figure 1. An example of the paired-associate material used by Cook and Kendler (1956)

In the prompting procedure the subject saw the stimulus term (see Figure 1), then the response term, and then responded overtly by drawing the code line. In the confirmation procedure the subject saw the stimulus term, then overtly responded, and finally saw the response term.

Cook (1961, 360) maintained that "the reinforcing effect of confirmation, whether early in learning or late, would simply overpower all other considerations." Kendler's view was that initially prompting would be superior to confirmation, but once the subject could spontaneously emit the response, as in later learning, confirmation would prove superior. The results indicated that neither prediction was correct; prompting was superior to confirmation both early and late in learning.

A theoretical model was invented which was not only compatible with the obtained differences in the number of correct responses, but which also yielded predictions for two other measures: (1) the number of different legitimate responses, and (2) the ratio of substitution errors to total errors. A legitimate response was defined as any of the ten responses used in the learning task, regardless of the stimulus to which it had been attached. A substitution error was defined as the evocation of a legitimate response by one of the stimuli other than the stimulus with which the response was paired.

The model (Figure 2) asserts that the learning of a single paired-associate item involves the learning of four S-R connections:

$$\begin{tikzcd}
r_1(s_1) \arrow[r] & r_2(s_2) \arrow[r] & R_R \\
S_1 \arrow[u] & S_2 \arrow[u] &
\end{tikzcd}$$

Figure 2. The Cook and Kendler (1956) model

(1) to the stimulus term (S_1), an implicit response (r_1) must be connected (alpha process), this implicit response (r_1) has certain stimulus properties (s_1); (2) the response term (S_2) must also be connected to an implicit response (r_2) (beta process) which has stimulus properties (s_2); (3) during training s_1 becomes connected to r_2 (gamma process); and (4) the stimulus s_2 cues off the overt response R (delta process).[1]

Two additional assumptions were made. Since S_1 was a familiar stimulus, it was assumed that the alpha process had already been established prior to the learning trials. The delta process was also assumed to be established since drawing a line did not involve a skill that had to be learned. Before a subject could respond correctly then, he would have had to first establish the beta connection, then establish the gamma connection. A correct response (R) to the stimulus (S_1) would then be possible through the mediational chain: $S_1 \longrightarrow r_1\ (s_1) \longrightarrow r_2\ (s_2) \longrightarrow R$.

The temporal priority of the beta connection resulted in the two predictions mentioned earlier: (1) anything that interferes with the beta process wil have its greatest effect early in learning, and (2) anything that interferes with the gamma process will have its greatest effect later in learning.

Analysis of the number of different legitimate responses showed no significant differences among the four groups. This finding indicated that the response terms as such were learned just as well under the confirmation procedure as under the prompting procedure. Since no interference with the beta process was found, prediction one could not be tested.

The model predicts that prompting will be superior to confirmation because: (1) the S-R delay between s_1 and r_2 is shorter in the prompting condition, and (2) in the confirmation procedure the subject is required to make an overt response which is usually incorrect early in learning. This incorrect response is, however, attached to s_1 and later interferes or competes with the $s_1 \rightarrow r_2$ hookup. Thus prompting results in a greater number of correct responses because it facilitates the gamma process and/or because the confirmation procedure interferes with it.

Having established that the gamma process was interfered with in the confirmation group, the authors proceeded to confirm the second prediction by showing that the ratio of substitution errors to total errors was not significantly different for the Prompt Early group and the Confirmation Early group but was significantly different for the Prompt Late versus the Confirmation Late group.

---

[1] Although the terminology alpha, beta, gamma, and delta did not appear until later, it is introduced at this point to avoid using two sets of notations for the same processes.

In summary then, the model presented was consistent with the observed differences in the number of correct responses an correctly predicts the results when the gamma process is interfered with.

That a prompting technique is superior to a confirmation technique was not a readily acceptable notion to some psychologists. While the results obtained by Cook and Kendler were not questioned, it was argued that training had not been continued long enough for the facilitative effect of confirmation (presumably the reinforcing effect) to exert its influence. The highest level of learning reached in the previous study was only 54.8 per cent. In addition, it was argued that in a confirmation procedure the subject makes an overt response without any cue support, which is what he must do on a test. This contention called into question the adequacy of the theoretical model which predicts the superiority of a prompting procedure.

In an attempt to test these opposing hypotheses (a prompting mode will be superior to a confirmation mode in the course of learning to mastery versus a confirmation mode will not be inferior to a prompting mode in the course of learning to mastery), as well as to verify the results of the Cook and Kendler study, Cook (1958) ran two groups of subjects to mastery. The task was the same as before: paired-associate learning of letters and code lines. Both the Prompted group and the Confirmation group received thirty-six training trials with a test after every third training trial (twelve tests in all). Analysis of the number of correct responses showed the Prompted group to be superior. (.05 level or above) to the Confirmation group on tests two through eight. Results on tests one and nine were in the hypothesized direction. Lack of significance on test one had been expected for the subjects were presumed to be still learning to establish a stable response (r_2) to the novel stimulus (S_2). On tests nine to twelve, the learning curves for both groups had reached an asymptote, or almost so, and the difference in the number of correct responses was held down by this ceiling. The results, then, did support the theoretical model as predicted.

Analysis of the number of different legitimate scores showed that the Prompted group learned significantly more responses than did the Confirmation group. This finding suggested that the prompting condition not only facilitated the gamma process, but also the beta process. While this finding was not consistent with the results of the Cook and Kendler (1956) study, it was not incompatible with the theoretical model or prediction number one from that study.

The author gives two possible reasons for the interference with the beta process in a confimation procedure.

> In the first place, in the early stages of learning under the confirmation procedure, the fact that S makes a wrong overt response just before the response term S_2 is presented, may cause the implicit response r_2 to become conditioned to the stimulus properties, which we shall call s_w, of that wrong response. This means that the $s_w - r_2$ connection is competing with the $S_2 - r_2$ connection to the detriment of the latter. In the second place, under the confirmation procedure when the code line S_2 flashes on the screen, S typically glances at it then looks at the code line he has just drawn to compare it with the correct one. This act, besides providing an interfering cue, reduces the effective duration of the stimulus S_2 and thus hinders the establishment of a stable implicit response to it (Cook, 1958, 10).[2]

---

[2] Cook, J. O. 1958. Processes underlying the learning of a single paired-associate item. This question is from a mimeographed full-length account of the study that appeared as a supplementary report in the J. Exp. Psychol., 1958.

Analysis of the substitution scores indicated that early in learning the prompted group made significantly more substitution errors, but that a reversal occurred later in learning and a significant difference was obtained in the opposite direction. This finding is consistent with the notion that a confirmation procedure interferes with: (1) the beta process allowing more r_2's tp be initially available for the Prompted group to substitute and (2) the gamma process thus providing the confirmation group with many free floating responses which are potential substitution errors later in learning.

In summary, the model would lead one to predict that, in terms of the number of correct responses, a Prompted group would prove superior to a Confirmation group. It would also lead to the prediction that early in learning, a prompting procedure should be superior to a confirmation procedure in terms of the number of different legitimate responses, and finally that in later learning a Confirmation group should have significantly more substitution errors than a Prompted group.

The reasons for the inferiority of the confirmation procedure are:

a. The S-R interval is longer than in a prompted procedure.
b. An overt response, usually wrong in early learning, occurs in the S-R interval.
c. The incorrect response may become attached to s_1 and offer competition to the correct $s_1 \longrightarrow r_2$ connection.
d. The wrong response has stimulus properties (s_w) which may become attached to r_2 and interfere with $S_2 \longrightarrow r_2$ connection.
e. In a confirmation procedure, the effective presentation time of S_2 is lowered because the subject typically compares his wrong overt response with the correct response.

Cook and Spitzer (1960) reasoned that if the subject was not required to make an overt response, then one source of interference in learning the mediational chain under a confirmation procedure would be eliminated. This change in procedure should have certain predictable consequences and these predictions could be tested. The predictions arrived at were: (1) that a confirmation procedure with overt responding should prove inferior to a confirmation with no overt responding group, and (2) a prompting procedure with no overt responding should prove superior to both types of confirmation and also a prompted overt responding group. By eliminating the overt practice in the confirmation group, the interfering effect of the incorrect response could be eliminated. Since the skill required to draw the code line has been learned prior to the experiment, overt practice in a prompted group merely detracts from the time available to practice the needed gamma process. Presumably what is practiced in this time interval in a prompted condition is not the beta process (since the stimulus S_2 is not present and the beta process cannot be practiced unless it is present) but rather the gamma process. Levine's (1963) data corroborate this notion.

Using the same materials as in the previous study, a 2 x 2 design, (prompting versus confirmation, overt versus no overt responding) was used to test the predictions. In terms of the number of correct responses, the results indicated that the Prompted No Overt Responding group was significantly superior to the Confirmation No Overt group as predicted, the Confirmation No Overt Responding group was significantly superior to the Confirmation Overt Responding group, the Prompted No Overt Responding group was superior to the Prompted Overt Responding group, and the Prompted Overt Responding group was superior to the Confirmation Overt Responding group as predicted. No predictions regarding the relative positions of the Prompted Overt Responding and Confirmation

No Overt Responding group were made because the relative importance of the two variables was unknown. Analysis of the number of different legitimate responses showed that both variables had a deleterious effect. The authors (p. 275) conclude that

> overt practice interferes both with learning the response terms as such and with connecting it with its proper stimulus and also that S-R delay interferes with the process of connecting a response to its proper stimulus but it has no very consistently reliable effect upon learning the response term as such.

Sidowski, Kopstein and Shillestad (1961) noted that in the Cook and Kendler (1956) study more efficient learning took place during a simultaneous presentation or "neutral" pretraining period than in the subsequent prompting trials. This finding is in agreement with that of Kopstein and Roshal (1955) who also reported findings which suggest that subjects learn more efficiently under simultaneous presentation of S-R terms. Using a paired associate task (English-Russian words) five experimental groups were compared with a Simultaneous Observation group (S-R terms presented together). The five groups were (1) Prompt (S term, then R term, then overt practice); (2) Prompt S-R (S term, then S-R term, then practice); (3) Confirmation (S term, then practice, then R term); (4) Confirmation-Prompt (S term, then practice, then S-R term, then practice); and (5) Confirmation S-R (S term, then practice, then S-R term).

Analysis of error data from test one (which was administered after nine practice trials) indicated that the Prompt S-R group and Simultaneous Observation group did not differ. The Prompt S-R group had significantly fewer errors than all other groups except the Simultaneous Observation group and that the Simultaneous Observation group had significantly fewer errors than the Confirmation, Confirmation S-R and Confirmation-Prompt group, but did not differ significantly from the Prompt group. No significant differences were obtained among the Confirmation, Confirmation S-R, Confirmation-Prompt and Prompt groups, although the Prompt group was in the expected direction of fewer errors.

Analysis of the error data from test two (administered after twelve practice trials and one test) indicated that the Prompt S-R group had significantly fewer errors than the Confirmation, Confirmation S-R and Confirmation-Prompt groups, but did not differ significantly from the Prompt and Simultaneous Observation group. The Simultaneous Observation group had sigificantly fewer errors than the Confirmation, Confirmation S-R, Confirmation-Prompt groups, and the Prompt group had significantly fewer errors than the Confirmation group. On a third test (R-S recall) the Simultaneous Observation and Prompt S-R groups had significantly fewer errors than the Confirmation group. No other differences were significant.

In no case, then, did the Simultaneous Observation group exceed the Prompt group by a significant amount, yet the results were all in the hypothesized direction. The Prompt group performed significantly better than the Confirmation group only on test two, although the results were all in the expected direction of a superiority for the prompted group.

It should be noted, however, that all groups except the Simultaneous Observation group practiced overtly. As Cook and Spitzer (1960) have shown, no overt responding facilitates the acquisition of the gamma process and the results of Sidowski et al. might be interpreted as a confirmation of the Cook and Spitzer study rather than showing a superiority of a simultaneous presentation of items.

Briggs (1961) has reported two studies in which prompting modes and confirmation modes of training were used. In the first study, four types of training procedures were compared: (1) Prompting (S term, then R term, then covert practice); (2) Single Try Confirmation (S term, then overt response, then confirmation of the response); (3) Multiple Try Confirmation (S term, as many overt responses as necessary to "hit" on the correct response, then confirmation); (4) Single Feedback (S term, then one response, then right or wrong confirmation). Three different problems, paired-associate learning, serial learning, and problem solving, were assigned to each practice mode. Each subject practiced on one task for twenty minutes. A test was administered after the training session and again two weeks later.

Overall analysis showed a significant difference between the tasks; the paired associate task was easiest and the problem solving task the most difficult. Significant differences between modes of practice were found on test one and two of the paired-associate task and on test one of the serial learning task.

Error scores at test one of the paired-associate task showed the Prompt mode to be significantly superior to all other methods and the Single Try Confirmation mode to be significantly superior to the Single Feedback mode. Analysis of the error data at test two of the paired-associate task showed the same results except that the difference between the Prompt mode and the Single Try Confirmation mode had dropped below the .05 level of significance.

The second study dealt with the relative effectiveness of three particular combinations of modes during practice. The same materials were used, and the subjects were again given twenty minutes training with an immediate test and another test two weeks later. The three practice modes were: (1) ten minutes prompting followed by ten minutes single try confirmation; (2) ten minutes prompting followed by ten minutes multiple try confirmation; and (3) five minutes prompting followed by five minutes single try confimation, then ten minutes multiple try confirmation.

Analysis of the error scores showed no significant differences among the training methods although group three showed a slight superiority over the other two groups.

On the basis of the first study, Briggs noted the general superiority of the Prompting mode, but, due to the lack of significant differences between modes on the problem solving task, he states that possibly the difficulty of learning the task is a factor in the relative training efficacy of the various modes.

It should be noted that under the prompting condition more practice trials were possible in the allotted twenty minutes since the subjects did not have to search for the correct response term. It would be possible then, that prompting was superior because more practice trials were used. No data on either the number of practice trials or the times taken per trial were reported and the effects of this variable cannot be evaluated.

Briggs attributes the lack of significance between groups in the second study as being due to too much practice under the prompting condition. He states (p. 385)

> since the prompting mode is an effective mode of learning, most of the mastery of the material was probably achieved before the second or third modes in the sequence were added.

He does feel, however, that a combination of prompting and confirmation trials might be superior to continued practice in the Prompting mode alone. He further states (p. 385) that "It would appear that optimum individual sequences could be worked out to fit the subject and material more precisely." Unfortunately the second study did not include a pure Prompted group for comparison with the mixed groups.

Taken together, the studies reviewed thus far seem to indicate that: (1) a mixed mode is better than a pure confirmation mode (Cook, 1958); (2) a pure prompting mode is superior to several different pure confirmation modes (Briggs, 1961); (3) there is probably no difference between some mixed modes (Briggs, 1961); and (4) it remains to be shown conclusively that a mixed mode is superior to a pure prompted mode.

Angell and Lumsdaine (1961a) differentiated between a prompting procedure and a confirmation procedure in terms of amount of cue support provided to the learner. In the prompting procedure, the response is produced with a maximum of cue support and in the confirmation procedure with a minimum of cue support.

They note that a "literal interpretation" of Cook's (1958) results might lead one to expect that a subject would learn fastest when 100 per cent cue support is provided throughout the learning process. None of the studies by Cook and his associates reviewed thus far had included a pure prompted group however, because every third prompting trial had been followed by a test trial.

Reasoning that in a mixed training sequence of both prompting and confirmation type trials the S̲ is afforded practice in responding to the unprompted situation which represents the learning criterion, Angell and Lumsdaine hypothesized that a sequence in which both prompting type and confirmation type training trials occur will be more effective than a pure prompted sequence.

To test this hypothesis, the same paired-associate learning task as used by Cook (1958) was learned under two modes of practice. All subjects received twenty-four practice trials and three test trials. In the "Pure" Prompt group, all practice trials were of the prompted type. A Mixed group received twenty-four practice trials where every fourth trial was a confirmation trial. A test was administered after twelve and twenty-four practice trials and another fifteen minutes after completion of test two.

Chi square values obtained at each of the three test points, using the number of persons with no errors, failed to reach the .05 level of significance (two-tailed probability). In each case the results were in the hypothesized direction.

When the proportion of persons making errors on test one, but no errors on test two, test three, or test two and test three was computed, significance was obtained only on the pooled date from test two and test three (one-tailed probability was .045).

The error data from test one was evaluated and a significant difference was found in the hypothesized direction (one-tailed probability was .047). When the total number of errors made by all subjects in both conditions was considered, the reported chi square median test value was just significant at the .05 level (two-tailed probability). There is some inconsistency in the figures reported however.

It should be noted that if the retention data (data from test three) is not included in the analysis, in only one instance (error data from test one) is significance reached. It would seem then, that the two modes of practice differed not so much on a measure of acquisition, but rather on a measure of retention. These results, while not conclusive, are interpreted as lending support to the stated hypothesis and the authors conclude that

> taken together, these results suggest the conclusion that a training sequence which incorporates some unprompted practice trials will be more effective than one which does not. The results of the Cook experiments, which showed the superiority of the prompting procedure, of course, are not questioned, because of differences in experimental procedures. Cook did not investigate a "pure" prompting condition, since unprompted trials were interspersed throughout his training sequence. On the other hand, the present experiment did not include a "pure" unprompted procedure such as Cook's confirmation condition. The results of Cook's experiments and the present one taken together indicate that while an unmixed prompting sequence is superior to an unmixed confirmation sequence, it is less effective than a procedure in which both types of trials occur (Angell and Lumsdaine, 1961, 397).

On the basis of their results, Angell and Lumsdaine further hypothesized that in learning a paired-associate task the optimum procedure would be a sequence in which prompting trials predominate in the early stages of learning with a gradual transition to more and more frequent confirmation trials.

This arrangement is similar to the vanishing technique recommended by Skinner (1958) in which the amount of stimulus support on successive presentations of the material is progressively and gradually diminished. This results in a gradual weaning of the learner from the initial stimulus support which is, according to Skinner, initially essential but terminally undesirable.

Cook et al. (1963) noted that one possible explanation for the superiority of the mixed training sequence over the pure prompted sequence might be that the mixed sequence provided knowledge of results to the learner. A pure prompted procedure provides no such information. To test this hypothesis, Ss practiced for fifteen prompted trials with tests interspersed after each third prompted trial. One group received tests with no feedback of results; a second group received tests with feedback of right or wrong; and a third group received tests with feedback of correct answers (i.e., a confirmation trial). No significant differences were found between the groups. In addition, these three mixed groups did not differ significantly from two pure prompted groups. These results seem to support in certain respects the finding of Angell and Lumsdaine (1961) who found no differences on a test given immediately after training.

The number of studies reported in the literature that deal with the problems of vanishing is not large, and it is frequently the case that divergent findings are reported. One possible explanation for these divergent findings may lie in the fact that there is no one method for vanishing a cue, and different methods may have different effects. Another difficulty lies in the fact that results of many of the studies done are not widely circulated. Due to the difficulty of obtaining full-length reports, many of the studies reviewed in the pages to come will, of necessity, be brief.

Kimble and Wulff (1961a) compared the performance of a Partially Guided (prompted) group with the performance of a No Guidance group in learning to read the C and D scales of a slide rule. Prompting consisted of telling the subjects the region on the scale in which the correct answer would fall. Analysis of the results indicated that on a test given immediately after training, the Partially Prompted group performed better than the No Guidance group. While a total Prompt group was not included, the authors state (p. 223) that in giving total support (_i.e._, an arrow pointing to the correct response)

> the learner would always make the correct responses, but the cues he would use might often be quite different from those he would have to use in the test situation and this kind of guidance might not be effective.

In the second study, Kimble and Wulff (1961b) compared the performance of a 100 per cent Guided group with the performance of a 50 per cent Guided group. In the 100 per cent Guided group, all sixteen practice examples had the region on the scale in which the correct answer would fall indicated. The 50 per cent Guided group had the region indicated on only eight of the practice examples. The 50 per cent Guided group did prove superior, but the results were not unconfounded. There was a lack of control for previous experience with the task (reading the C and D scale of a slide rule), and ability, which had been shown to be related to performance, was not stable from block to block.

Macoby and Sheffield (1961) have shown that subjects who practiced initially after short segments of a demonstration film, then after longer segments on the next trial, and finally after seeing the demonstration as a whole, maximized transfer from training conditions to conditions of criterion performance without demonstrational support. This "weaning" method was superior not only to a method that provided less direct prompting initially, but also to a method that continued to use short step prompting.

These three studies (Kimble and Wulff, 1961a, 1961b) and Macoby and Sheffield (1961) lend support to Lumsdaine's hypothesis.

Angell and Lumsdaine (1962) used a fifty-four frame program to teach a short-cut method for squaring numbers ending in "5" to fifth and sixth graders. Two training procedures were compared: (1) a version of the program that employed vanishing, and (2) a non-vanishing version in which full prompting was continued throughout the program. Analysis of scores on a test administered immediately after training failed to show a significant difference between the two versions. Significance was obtained, however, on a test given two weeks later where the Vanished group out-performed the Non-vanished group. The fact that effects of a theoretically important factor showed up only after a lapse of time suggests that retention tests may be needed for proper evaluation of programs.

Stolurow and Lippert (1963) used a 900-frame program to teach twenty-five nouns to educable mentally handicapped school children. The effects of using varying numbers of prompting trials followed by various numbers of confirmation trials on the learning and retention of the nouns was studied. Results indicated that the median number of errors was zero under each sequence and that retention was maintained at about the same level after one, seven, or thirty days. The effects of vanishing the prompt by introducing the confirmation sequence after various numbers of prompting trials was complex and not entirely consistent.

Results of a second study by Stolurow (1961) indicated that the relative number of prompted and confirmation trials did not seem to be as important in predicting learning as the total number of trials regardless of the balance.

Another finding of no difference was reported by Campbell (1961). Using a linear program on static electricity, two versions of the program were compared:

(1) a rather fully prompted version, and (2) a version with more indirect cueing. Small nonsignificant differences of the order of 2 per cent on a criterion test were found.

In a study designed to test the effectiveness of a partial cueing method, Angell and Lumsdaine (1961) compared the learning of a paired-associate task under a standard confirmation technique with learning under a partial cueing method. In the partial cueing method, the subject could see one, two, or three letters of a three-letter symbol which constituted the response term. Slow learners learned somewhat more of the hard items under the partial cueing method while fast learners did not learn as many of the easier items under this method. Overall differences were not significant.

Rigney and Budnoff (1962) hypothesized that a prompting-confirmation procedure would be superior to a confirmation-confirmation combination for the acquisition of associations and to a prompting-prompting combination for the retention of associations. Using a 107-frame program on Boolean algebra, these predictions were tested using 105 high school students as subjects. Neither of the hypotheses was confirmed; tests showed that the confirmation-confirmation treatment resulted in lower error scores on the acquisition tests than did the prompting-confirmation treatment, and that these effects decreased over retention intervals.

Guthrie and Lumsdaine (1961) compared performance on a paired-associate learning task under three conditions of training. One group received a confirmation procedure; another a partial prompting (S term, then a weakened exposure of the R term) procedure; and third, a fully prompted group. The Partial Prompted group was significantly superior to the Confirmation group, but the performance of the Fully Prompted group led the authors to state (p. 17) that

> immediate full prompting may frequently represent an optimum condition and that the various partial cueing methods simply provided varying approximations to this immediate, full prompt condition.

They conclude that the conditions under which partial prompting is superior to full prompting are not yet fully understood.

In summary, it can be said that the studies using a paired-associate task (Cook and Kendler, 1956; Cook, 1958; Cook and Spitzer, 1960; Guthrie and Lumsdaine, 1961; Briggs, 1961; and Kopstein et al., 1961) have shown that a mised mode of practice trials is superior to a pure confirmation mode. The issue of whether a mixed mode is superior to a pure prompting mode does not seem to be cleanly settled. Angell and Lumsdaine (1961) and Cook et al. (1962) reported somewhat different results. Briggs (1961), however, has shown that several different mixtures do not differ significantly. Findings of no difference were reported by Angell and Lumsdaine (1961b) between a confirmation mode and a partial cueing (prompting) mode.

Of the studies using tasks other than paired-associate learning, two, Kimble and Wulff (1961a) and Macoby and Sheffield (1961), can be interpreted as supporting the findings of a superiority in favor of the mixed mode over the pure confirmation mode. Angell and Lumsdaine (1962), Kimble and Wulff (1961b) and Macoby and Sheffield (1961) present data which support the conclusion that a mixed mode is superior to a pure prompted mode.

The study by Stolurow and Lippert (1963) supported the finding of Briggs (1961) of no difference between mixed modes. Lastly, studies by Rigney and Budnoff (1962) and Campbell (1961) reported a finding of no significant difference between pure confirmation, pure prompting and mixed modes.

Experiment A

Statement of Problem

The objective of the present study was twofold. First, it was proposed that the hypothesis which states that a sequence in which both prompting and confirmation type trials occur will be more effective than a pure prompted sequence (Angell and Lumsdaine, 1961a) be tested against an alternative hypothesis based on the Cook and Kendler model. This seems desirable for several reasons. Angell and Lumsdaine did not use a pure prompting condition since a test was scheduled between the twelfth and thirteenth prompting trials. In a strict sense, their hypothesis remains untested. In addition, the analysis at the testing points during and immediately after training failed to yield consistently reliable differences. Significance was obtained only after data from a test given slightly more than fifteen minutes after the last training trial was included. A question arises as to whether this data can legitimately be used to evaluate their hypothesis. This confusion results from the lack of an operational definition for the term "effective." If their interpretation of a "more effective technique" included a superiority of one method over another using some retention measure, then the use of such data cannot be questioned, and the hypothesis can be accepted. If, on the other hand, we interpret "effective" to mean a superiority as reflected by some acquisition ("short term" retention) measure, acceptance of the hypothesis is risky. A third possibility is that they intended the term "effective" to cover both acquisition and retention measures. As noted earlier the effects of a vanishing technique may not make themselves felt on a test of acquisition but may influence results obtained on a test of retention. If this is the case, it would seem that an operational definition of "effective" is of theoretical importance.

Angell and Lumsdaine tested their hypothesis using a three to one ratio of prompting to confirmation trials arranged in a specific sequence. It was felt that a replication of their findings should be attempted, but using other ratios and other combinations of trials.

The phrase "more effective" will be operationally defined as indicating a difference between groups, significant at the .05 level or better, on a measure of the number of correct responses emitted or on a measure of the number of correct response terms, as such, that have been learned.

The second objective of the present study was to test the hypothesis which states that

> In a paired-associate task, the optimum procedure is a sequence in which prompting trials predominate in the early stages of learning with a gradual transition to more and more frequent confirmation trials (Angell and Lumsdaine, 1961a, 397)

against four alternative schedules. These alternatives would include: (1) confirmation type trials predominating early in learning with a gradual transition to more and more frequent prompting trials (CE); (2) regularly-spaced confirmation trials (RS); (3) no confirmation trials at all (pure prompt, PP); and (4) letting each subject decide when and how many confirmation trials he wants with a mandatory test at the end of the training sequence (SD).

Angell and Lumsdaine predicted that the Prompt Early (PE) group will be superior to the Confirmation Early (CE) group and Pure Prompt (PP) group. No predictions about the relative positions of the RS and SD groups were made. The Cook and Kendler model would predict the PP group to be superior to all others except the SD group. No predictions are made as to the relative position of the SD group for it is not known what ratio of prompting to confirmation trials will be chosen. Nor is it known what the criteria of student choices are. A prediction based on the Cook and Kendler model would be that the closer to the PP group their choices become, the better will be their relative position. No predictions are made concerning the positions of the CE, PE, and the RS group, for while the proportion of prompt to confirmation trials will remain constant (50-50), the relative effects of proactive and retroactive inhibition are not known.

Method

Materials.- The paired-associate materials were the same as those used by Cook and Kendler (1956), Cook (1958), and Cook and Spitzer (1960). The S_1 term of each pair was one of the first ten letters of the alphabet and the S_2 term was a bent line that connected three dots in a field of seven randomly placed dots.

Three film-strip projectors were used to present the material on a screen. The timing of the intervals and duration of S_1 and S_2 terms is shown in Figure 3, and was done automatically. Tones of 300 cycles per second were programmed on a tape which, when run through a tape recorder, activated a switching relay and advanced the film. Three films were used: one consisted of twelve confirmation trials, one consisted of twelve prompting trials, and a third contained the test. It was possible to select any twelve-trial combination of prompting and confirmation trials by merely throwing a switch from prompt to confirmation. A thirty-second intertrial interval provided sufficient time for this change.

PROMPT	.5 sec view letter	2 sec view code line	3 sec draw code line	1 sec interval
CONFIRMATION	.5 sec view letter	3 sec draw code line	2 sec view code line	1 sec interval
TEST	.5 sec view letter	5 sec draw code line		1 sec interval

Figure 3. Timing of intervals and duration of S_1 and S_2 terms

Design.- A 5 x 4 factorial design utilizing twenty groups was employed. Five different sequences of training trials (prompt trials early in training, confirmation trials early in training, regularly spaced prompted trials, pure prompted training, and a sequence in which Ss were allowed to choose either a prompting or confirmation method prior to each practice trial) were compared over four levels of training (three, six, nine or twelve training trials). The specific arrangement of prompt and confirmation trials for each group is shown in Figure 4.

Number of training trials	Prompt Early	Conf. Early	Regular Space	Pure Prompt	Self Determined
3	PPP T	CCC T	PCP T	PPP T	--- T
6	PPPPCP T	CCCCPC T	PCPCPC T	PPPPPP T	------ T
9	PPPPCPCPC T	CCCCPCPCP T	PCPCPCPCP T	PPPPPPPPP T	--------- T
12	PPPPCPCPCCCC T	CCCCPCPCPPPP T	PCPCPCPCPCPC T	PPPPPPPPPPPP T	------------ T

Figure 4. The specific arrangement of prompt and confirmation trials for each group

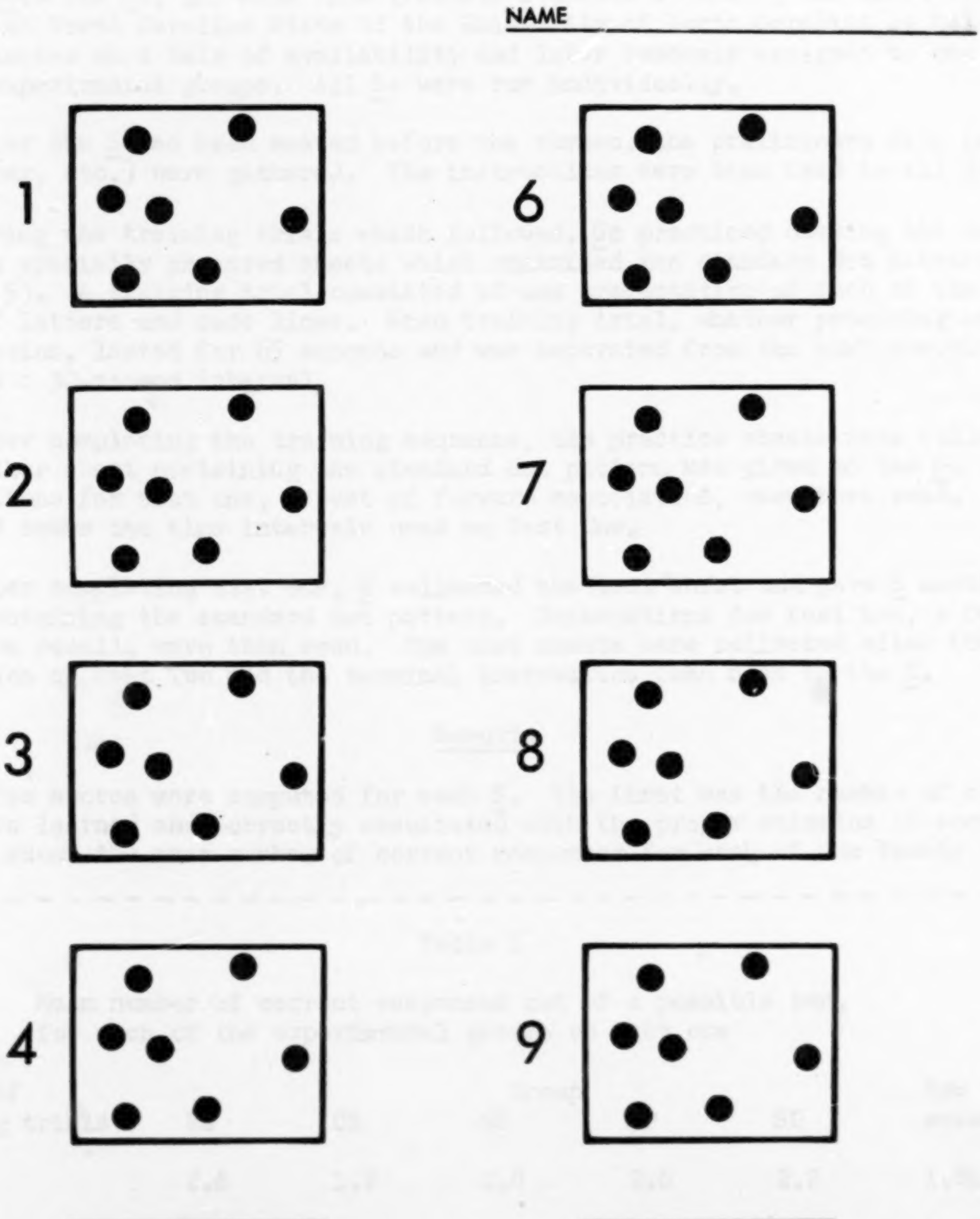

NAME
1
2
3
4
5
6
7
8
9
10

Figure 5. The response sheet.

Procedure.- The *Ss*, 100 male undergraduate students attending the first summer session at North Carolina State of the University of North Carolina at Raleigh, were selected on a bais of availability and later randomly assigned to one of twenty experimental groups. All *Ss* were run individually.

After the *S* had been seated before the screen, the preliminary data (name, IBM number, etc.) were gathered. The instructions were then read to all *Ss*.

During the training trials which followed, *Ss* practiced drawing the code lines on specially prepared sheets which contained ten standard dot patterns (Figure 5). A training trial consisted of one presentation of each of the ten pairs of letters and code lines. Each training trial, whether prompting or confirmation, lasted for 65 seconds and was separated from the next practice trial by a 30-second interval.

After completing the training sequence, the practice sheets were collected and another sheet containing the standard dot pattern was given to the *S*. Instructions for test one, a test of forward association, were then read. Figure 3 shows the time intervals used on Test One.

After completing test one, *E* collected the test sheet and gave *S* another sheet containing the standard dot pattern. Instructions for test two, a test of R-term recall, were then read. The test sheets were collected after the completion of Test Two and the terminal instruction then read to the *S*.

Results

Three scores were computed for each *S*. The first was the number of responses that were learned and correctly associated with the proper stimulus (C score). Table 1 shows the mean number of correct responses for each of the twenty groups.

Table 1

Mean number of correct responses out of a possible ten, for each of the experimental groups on test one

Number of training trials	Group FE	CE	RS	PP	SD	Row means
3	2.6	1.2	0.8	2.6	2.2	1.84
6	6.2	3.0	3.6	4.8	4.2	4.36
9	5.2	6.0	6.4	4.0	7.6	5.88
12	4.6	5.2	5.8	6.8	6.6	5.80
Column means	4.65	3.85	4.10	4.55	5.15	

The second score computed was the number of different legitimate responses elicited on test one (D score). A ligitimate response is any one of the ten code lines, whether elecited by the correct stimulus or not. Duplicated legitimate responses on the same test were ignored. Table 2 shows the mean number of different legitimate responses elicited on test one.

Table 2

Mean number of different legitimate responses, out of a possible ten, for the experimental groups elecited on test one

Number of training trials	Group PE	CE	RS	PP	SD	Row means
3	4.0	3.4	3.0	3.4	4.4	3.64
6	7.8	4.6	5.4	5.8	5.0	5.72
9	6.0	6.6	7.2	5.0	7.8	6.52
12	5.2	6.2	6.4	7.2	7.4	6.52
Column means	5.75	5.25	5.50	5.25	6.15	

The third score computed was the number of different legitimate responses emitted on test two. These data are presented in Table 3.

Table 3

Mean number of different legitimate responses, out of a possible ten, emitted on test two

Number of training trials	Group PE	CE	RS	PP	SD	Row means
3	4.2	5.4	3.8	5.6	5.0	4.80
6	8.2	5.2	6.2	5.8	6.0	6.28
9	6.3	7.0	7.8	6.0	9.0	7.20
12	6.8	7.2	7.6	7.8	7.2	7.32
Column means	6.35	6.20	6.35	6.30	6.80	

Since the data were not normally distributed (X^2 with 7 d. f. was 21.5), a generalisation of the Kruskal-Wallis ANOVA (Drewes and Bishir, 1963) was used to test for significance among the groups.

Drewes has shown that the H statistic computed over cells is a linear sum of information provided by row, column, and treatment effects. The informational components have been shown to be distributed as X^2 subject to the assumptions of asymptotic normality and a constant number of replications over cells.

The data for the five replications of each of the twenty groups was combined and then ranked. Sums of ranks for twenty groups were obtained and an *H* computed. This value appears in tables as the Total *H*. Sums of ranks down columns were obtained for five groups and an *H* computed. This value appears in the tables as the Group *H*. Sums of ranks across rows were obtained for four levels of training and an *H* computed. This value appears in the tables as the Training *H*.

Using the C scores, a total *H* value of 46.706 (corrected for ties) was obtained. Table 4 shows that the groups *H* was not significant nor was the interaction *H*. The value of *H* for the level of training variable was highly significant, indicating that *Ss* learned. It should be noted that tables of means are presented but the analysis is done using ranks.

Table 4

Results of the analysis of variance using C scores

Source	d.f.	H	P
Total	19	46.706	.002
Groups	4	2.874	$.3 > p > .5$
Training	3	38.076	.001
Interaction	12	5.783	$.9 > p > .95$

Analysis of the D scores from test one (Table 5) yielded the same results; the total *H* was highly significant, but neither the group *H* nor the interaction *H* reached or exceeded the .05 level of significance. The *H* value for level of training proved to be the major contributor to the obtained total *H* value.

Rank sum multiple comparisons (Steel, 1961) were done for each of the five main groups at each training level. In no instance did a significant improvement in C scores occur between three and six training trials. A significant improvement was found between three and nine training trials for all five groups and a significant improvement was found between three and twelve training trials for the CE and SD groups. No other comparisons proved significant.

Table 5

Results of the analysis of variance using D scores

Source	d.f.	H	P
Total	19	41.714	.01
Groups	4	1.993	.7 > p > .8
Training	3	27.106	.01
Interaction	12	12.705	.3 > p > .5

Discussion

The results of the present experiment cannot be interpreted as supporting the hypothesis which states that some sequences in which both prompting and confirmation type trials occur will be more effective than a pure prompted sequence. Although Angell and Lumsdaine (1961a) accepted a stronger version of this hypothesis, it was noted that their conclusions were based on retention as well as acquisition data. The present study, which was limited to acquisition data, does agree in part with the Angell and Lumsdaine study however. Limiting their analysis to acquisition data, Angell and Lumsdaine found significance in only one instance, and here the one-tailed probability was only .0475. The present results are also in agreement with the findings of Angell and Lumsdaine (1962) who reported a finding of no difference between a Fully Prompted group and a group which had cues gradually reduced (vanished).

Having established to their satisfaction that a mixed sequence of prompting and confirmation trials was superior to a pure prompting sequence, Angell and Lumsdaine (1961a) further hypothesized that a mixed sequence in which prompting trials predominate early in learning would be superior to other combinations. This hypothesis is also advanced by Skinner (1958, 1959) and is based on the notion that cue support must be lessened in order to insure that the learner is forced to obtain practice in responding with no prompt and thus does not continue to depend on the prompt.

Skinner seems to imply that this is a general rule applicable to all programs in which prompting is used. But it may be argued that this rule cannot be legitimately generalized to all situations. In a typical "Skinnerian" frame, a semantic variation may be used as a cue or an obvious transposition may be used. Figure 6 is an example of the latter.

Gross profit less overhead equals net profit; so

to get net profit, you subtract _____ from _____.

Figure 6. The use of an obvious transposition as a cue.

In this case, the cue support is always present and the possibility does exist that the learner will continue to depend on this support for correct answers.

For material programmed differently, this cue support is not always available. For example, in the prompted sequences of the present study (letter, then code line, then draw the code line), the learner responds (draws the code line) in the absence of the prompt; he does not respond until *after* the screen has gone blank. Presumably he is responding to the stimulus property (s_2) of the implicit response (r_2) to the external cue (S_2), and this is just what he *must* do on a test. It will be recalled that a correct response is made through the mediational chain $S_1 \rightarrow r_1(s_1) \longrightarrow r_2(s_2) \longrightarrow R$. The prompt merely insures that this implicit response occurs in close temporal contiguity with: (1) the stimulus to which it must become attached, and (2) with the response which it must elicit. It would seem then, that whether or not a prompt should be vanished depends on the way in which the material is presented.

It should be noted that while the Cook and Kendler (1956) model was invented to handle a prompting procedure like that employed in the present study, it could also be generalized to a situation in which the prompt was always available. In this situation, the mediational chain learned (if no vanishing was used) might be $S_1 \longrightarrow r_1 \longrightarrow S_2 \longrightarrow R$. The model would predict that in this instance performance would decline on a test where the S_2 was not presented, for the mediational chain would be incomplete. Skinner's "general" rule cannot be generalized to suit the occasion.

This analysis has said nothing about the possible motivational effects of a confirmation procedure. It might be hypothesized that a confirmation procedure, while not embodying the accepted conditioning principle as does the prompting procedure, might have a positive effect because it motivated the learner. It might also be used as a cue to tell the learner which items he does not know, thus allowing him to concentrate his efforts on learning them. We have already seen, however, that Cook *et al*., (1963) found no difference between a pure prompting group and a group that had a test after each set of three prompting trials. In addition to supporting the present findings, the Cook *et al*. study seems to rule out the possible superiority of a mixed sequence due to either motivation or knowledge of results. It could be argued that a replication of their study, using *Ss* from the general population would be useful. It is possible that college students were already highly motivated when performing in this experiment and thus the facilitative effect of a confirmation trial (due to its increasing motivation) was not experienced.

The results of the present study did not show a clearcut superiority for the pure prompted group. One possible explanation for this finding is suggested by the studies by Briggs (1961). It will be recalled that Briggs failed to find a significant difference between groups that were given initial practice using a prompting mode followed by various types of confirmation modes. This finding was especially surprising because that author had previously shown that significant differences existed between these confirmation modes. It was suggested that too much practice had been allowed under the prompting mode. The efficiency of the prompted practice outweighed all other considerations. Briggs suggested that in order to test more effectively for differences among mode combinations, the prompting mode "would have to be reduced to a minimum." (Briggs, 1961, 385).

Another finding that is relevant to the present study was the fact that decreasing the amount of prompted practice from ten to five minutes had no effect on the amount that was learned. This result supports the findings of the present study; the PP group (twelve prompted trials) did not differ from groups that received half as many prompted trials. It would seem that what is important is not how many trials are of the prompting type, but rather are there any prompting trials.

One would hypothesize, however, that there would be a limit below which this generalization would not hold. The present study has shown only that six prompting trials are about as effective as twelve. Future research might be directed toward discovering this limit, using both number of prompted trials and time as the independent variable. The interacting effects of amount of prompted practice and ratio of prompting and confirmation trials might also be investigated.

Analysis of Table 1 supports this explanation. Note the poor performance of group CE on trials one to six where only one prompting trial is given. On trials seven to nine, performance improved significantly, presumably due to the occurrence of the prompting trials. An initial deficit was overcome by a few prompting trials so that performance exceeded that of an Early Prompted group and a Pure Prompted group.

This explanation is also in agreement with the findings of Scarf (1961) and Hescart (1961) who found that when frames are highly prompted, the presence or absence of confirmation of responses had no effect on the amount learned.

Another explanation for the finding of no significance differences among groups might be that the n used in the present study was not large enough. Using the best performing group (group SD) and the poorest group (group CE), the n needed to obtain a significant difference (.05 level) was calculated. A Beta level of .20 and a difference of .5 resulted in an obtained value of 746.

A third possible explanation was suggested by the work of Angell and Lumsdaine (1962) and also that by Jacobs (1962). Jacobs noted that while the objectives of programming instructional material differ from the objectives of test construction (tests are used to predict or define the criterion behaviors while programs are used to modify criterion behaviors), there are similarities between the construction of each.

One of these similarities is the specification of some criterion behavior(s) in operational terms. As is often the case, however, measures of the ultimate criterion (what the test constructor is really interested in) are not readily available. Proximate criteria (what, for any number of reasons, the test constructor is willing to settle for) are often used as substitutes.

An example of a proximate criterion might be the use of the number of errors made while working through the program as the dependent variable. As Schramm (1962) notes, one of the presumed characteristics of an efficient program was that the learner made few errors as he responded to each item in the program. Both Stevens (1950) and Fry (1960) have shown, however, that a proximate criterion of errors during training did not reflect their ultimate criterion of performance on a post-test. The present study also supports those results.

Angell and Lumsdaine (1962) have shown that differences between a Vanished and No Vanished group did not appear until a retention test was given.

If we assume that an ultimate criterion for evaluating the effectiveness of several methods of programming material would include both measures of acquisition and retention, we have a third possible explanation for the results of the present study. It would seem that a test of long term retention should have been included as a dependent variable. This was not done since the semester ended just after the acquisition data was collected and many of the Ss were no longer available.

This analysis indicates that the proximate criterion, number of correct responses on a test given immediately after training, may not reflect the true differences among combinations of modes of practice. A re-evaluation of prior findings using a measure of retention as the dependent variable would seem to be in order.

Of particular interest is the way in which Ss in the SD group programmed the sequences of training trials. It was this group that was allowed to choose between a prompting and confirmation method before each training trial. Figure 7 shows the choices for each of the twenty Ss in this group.

Number of training trials	Practice trial											
	1	2	3	4	5	6	7	8	9	10	11	12
3	P	P	P									
	P	P	P									
	P	P	C									
	P	P	C									
	P	P	C									
6	P	P	P	P	C	P						
	P	P	C	P	C	P						
	P	P	P	C	C	C						
	P	P	C	P	P	P						
	P	P	P	C	P	P						
9	P	P	C	P	P	P	P	P	P			
	P	P	C	P	P	C	P	C	C			
	P	P	P	P	P	P	P	P	P			
	P	P	P	P	C	P	C	P	C			
	P	P	P	C	P	P	C	P	C			
12	P	P	P	P	C	P	P	C	P	C	P	C
	P	P	P	P	C	P	P	C	P	C	P	C
	P	P	P	P	P	P	P	P	P	P	C	C
	P	P	C	P	C	P	C	P	C	P	C	P
	P	C	P	P	P	P	P	P	P	P	C	C

Figure 7. Sequencing of prompting and confirmation trials by the three levels of the SD group

---

Generally, it can be said that this group chose a program that closely approximated the PE group. As Figure 5 shows, the ratio of prompted trials to confirmation trials remained fairly constant across the subgroups. For example, the ratio of the first three trials were 12:3, 13:2, 13:2, and 13:2 for subgroups 3, 6, 9, and 12, respectively. There was also a notable decline in the number of prompting trials chosen as practice continued. For example, the subgroup receiving twelve practice trials chose a 1:1 ratio on the last set of three trials.

Post-experimental inquiry indicated that the Ss in this group felt a prompting trial would be the better choice because they could learn under a prompting trial but not a confirmation trial. As one S put it, "Prompting sounded the best because it followed closer" (code line followed the letter), or as another S said, "I figured that prompting was the best way to learn it. Figured I could see the code line twice" (once when the prompt was presented, and again after he had drawn it) "and that way have more time to learn it." In every instance, the explanation given for taking a confirmation trial was to test how much had been learned. Usually Ss returned to prompting trials after the confirmation trial to "try to pick up what I didn't know." They expressed the feeling that learning would not occur using a confirmation procedure.

In an attempt to validate the assumption that the number of different legitimate responses learned is accurately reflected in the number of such responses evoked on test one, a test of R term recall was given. It was possible that since test one was not self paced, and the task of the subject was not merely to list the code lines that he could recall but to draw a specific code line to a specific stimulus, that the D score computed was not an accurate reflection of what the S had really learned. By giving the S more time and by instructing him to just draw as many of the code lines that he could recall, it was thought that a more accurate measure could be obtained. As expected, the D score measured in this manner was slightly higher than the D score computed from test one, but it did not differ significantly from it. The assumption that D scores are accurately measured on an automatically paced forward association test (as was test one) was not disproved.

Experiment B

Statement of Problem

In the previous experiment, the finding of no significant differences among groups resulted in no logically defensible basis for making conclusions about the true state of affairs. As Fisher (1942) has noted, the null hypothesis can only be disproved, it cannot be proved. It was felt that if a Pure Confirmation group had been included in the design, several hypothesized reasons for this finding of no difference might have been evaluated.

The notion was to use this Pure Confirmation group as a bench mark. Previous studies are almost unanimous in their findings that a mixed sequence results in significantly superior performance when compared to the performance of a Pure Confirmation group. If this finding could not be replicated, then it would seem that some experimental manipulation, difference in subject population, or experimenter bias had influenced the outcome. If, however, the five groups which had practiced with prompting trials included in the training sequence had all proved to be superior to the Pure Confirmation group, then the conclusion that the true state of affairs was no difference between these five groups might appear to be a more attractive hypothesis.

Method

Materials.- The materials were the same as used in Experiment A of this study. The timing of intervals and duration of S_1 and S_2 terms remained the same.

Design.- The experimental design consisted of four main groups. In each group, subjects received twelve training trials with a test following the last practice trial. Groups PE, CE, and PP were the same as in Experiment A. Group PC (pure confirmation) was the only new group.

Procedure.- The Ss, 45 male undergraduate students attending the second summer session at North Carolina State of the University of North Carolina at Raleigh, were selected on a basis of availability and later randomly assigned to one of the four groups. Ten Ss were run in groups PE, CE, and PP, and fifteen Ss were run in group PC. the procedure was the same as in Experiment A.

Results

Since the data were not normally distributed, the Kruskal-Wallis ANOVA was used. First, the ten Ss from groups PE, CE, and PP and ten Ss randomly drawn from group PC were compared. The obtained H value did not approach the .1 level of significance. Next, the data from the Ss in Experiment A at training level twelve (groups PE, CE, and PP) was added to the data from the present study. This resulted in an n of fifteen for each group. The obtained H value again failed to reach the .1 level of significance.

The means and standard deviation of the four groups are shown in Table 6.

Table 6

Means and standard deviations of the four experimental groups

	GROUP							
	PE		CE		PP		PC	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Experiment A	4.6	3.29	5.2	2.28	6.8	1.92	---	----
Experiment B	5.8	3.65	6.7	2.50	6.9	2.79	7.1	2.43

Discussion

The finding of no significant differences between groups PE, CE, and PP confirms the findings of Experiment A. That these three groups did not differ from the PC group is surprising. Previous findings by Cook and Kendler (1956), Cook (1958), Cook and Spitzer (1960), Angell and Lumsdaine (1961), and Briggs (1961) have shown this difference to be a reliable one. Indeed, the results of the present study were in the opposite direction, but not significant, the PC group exceeded the PP group. This finding explains the results of Experiment A. If performance after prompted training does not differ from that obtained after confirmation training, then different sequences of prompting and confirmation trials could not be expected to differ.

The present finding of no significant differences among the groups must be added to the growing number of studies on programmed instruction which do not succeed in disproving the null hypothesis. As Schramm (1962) notes (p. 4), these studies

> may indeed, be proving that no significant difference exists, but the suspicion arises that in many cases the programs are too short, the samples too small, the measuring instruments too dull, to pick up the differences if they exist.

Summary and Conclusions

This study attempted to test two alternate predictions concerning the effects of sequencing prompting and confirmation trials. According to Skinner (1958) and Angell and Lumsdaine (1961), performance after practice using a mixture of prompting and confirmation trials should be superior to performance after practice using a pure prompting procedure. A theoretical model presented by Cook and Kendler (1956) would lead one to predict the opposite outcome. Five experimental groups (mixed trials with prompting trials massed early in training, mixed trials with confirmation trials massed early in training, regularly spaced prompting and confirmation trials, a pure prompt sequence and a group that was allowed to choose either mode prior to each practice trial) were compared over four levels of training. Analysis of the C and D scores indicated that the treatments did not differ significantly.

References

Angell, D. and Lumsdaine, A. A. 1961a. Prompted and unprompted trials versus prompted trials only in paired-associate learning, pp. 389-398. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction*. National Academy of Sciences - National Research Council, Washington, D. C.

Angell, D. and Lumsdaine, A. A. 1961b. *A study of subject controlled partial cueing in paired-associate learning.* American Institute for Research, San Mateo, California. AFOSR- TN-61-1342.

Angell, D. and Lumsdaine, A. A. 1962. *Retention of material presented by auto-instructional programs which vanish and which do not vanish verbal cues.* American Institute for Research, Palo Alto, California. AFOSR-TN-62-1342.

Briggs, L. J. 1961. Prompting and confirmation conditions for three learning tasks employing the subject matter trainer, pp. 375-387. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction.* National Academy of Sciences - National Research Council, Washington, D. C.

Campbell, V. N. 1961. *Adjusting self-instruction programs to individual differences; studies of cueing, responding, and bypassing.* American Institute for Research, San Mateo, California. AIR-C41-5-62-FR.

Cook, J. O. 1958. Supplementary report: Processes underlying the learning of a single paired-associate item. *J. Exp. Psychol.*, 56, 455.

Cook, J. O. 1961. From audience participation to paired-associate learning, pp 351-366. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction*. National Academy of Sciences - National Research Council, Washington, D. C.

Cook, J. O. and Kendler, T. S. 1956. A theoretical model to explain some paired-associate learning data, pp 90-98. In G. Finch and F. Cameron (eds.), *Symposium on Air Force Human Engineering, Personnel and Training Research*. National Academy of Sciences - National Research Council, Washington, D. C.

Cook, J. O., Miller, N. G., Grier, J. B., and Stenan, J. W. 1963. Prompting and Confirmation in paired-associate learning. In Miller, N. G. and Cook, J. O. (eds.), *Studies in guided learning*. Report to U. S. Office of Education, 1963.

Cook, J. O. and Spitzer, M. E. 1960. Supplementary report: prompting versus confirmation in paired-associate learning. *J. Exp. Psychol.*, 59, 257-259.

Drewes, D. W. and Bishir, J. 1963. *A non-parametric analysis of variance for higher order designs.* Unpublished paper. Department of Psychology, North Carolina State of the University of North Carolina at Raleigh.

Fisher, R. A. 1942. *The Design of Experiments.* (3rd ed.) Edinburg: Olive and Boyd.

Fry, E. B. 1960. *Teaching Machines: An investigation of constructed response versus multiple choice methods of response.* Unpublished Doctoral Dissertation, University of Southern California.

Guthrie, P. M. and Lumsdaine, A. A. 1961. *Some effects of gradual partial cueing on the learning of paired associates.* American Institute for Research, San Mateo, California. AFOSR-TN-61-1341.

Hessert, R. B. 1961. *A comparison of three methods of programming study materials for auto-instruction.* Unpublished Doctoral Dissertation, University of Pittsburgh.

Jacobs, P. J. 1962. *Some implications of testing procedures for auto-instructional programming.* Aerospace Medical Division, 657th. Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio. Report No. MLR-TDR-62-67.

Kendler, T. S., Cook, J. O. and Kendler, H. H. 1956. *An investigation of the interacting effects of retention and audience participation on learning from films.* Air Force Personnel Training Research Center, Training Aids Research Laboratory, Chanute Air Force Base, Illinois. Lab. Note TARL-LN-56-2.

Kendler, T. S., Kendler, H. H. and Cook, J. O. 1954. *Effect of opportunity and instructions to practice during a training film on initial recall and retention.* Air Force Personnel Training Center, Training Aids Research Laboratory, Chanute Air Force Base, Illinois. Staff Research Memo.

Kimble, G. A. and Wulff, J. J. 1961a. "Response guidance" as a factor in the value of student participation in film instruction. pp 217-226. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction.* National Academy of Sciences - National Research Council, Washington, D. C.

Kimble, G. A. and Wulff, J. J. 1961b. The effectiveness of instruction in reading a scale as influenced by the relative amounts of demonstration and problem solving practice. pp 227-239. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction.* National Academy of Sciences - National Research Council, Washington, D. C.

Kopstein, F. F. and Roshal, S. M. 1955. Methods of presenting word pairs as a factor in foreign vocabulary learning. *Am. Psychol.*, 10, 354.

Levine, J. M. 1963. *Prompting versus confirmation in paired-associate learning under conditions of familiarity and novelty of stimulus and response terms.* Unpublished Master's thesis, Department of Psychology, North Carolina State of the University of North Carolina at Raleigh.

Lumsdaine, A. A. 1960. Some issues concerning devices and programs for automated learning. pp 517-539. In A. A. Lumsdaine and R. Glaser (eds.), *Teaching Machines and Programmed Learning, A Source Book*. National Education Association of the United States, Washington, D. C.

Macoby, N. and Sheffield, F. O. 1961. Combining practice with demonstration in teaching complex sequences: Summary and interpretation. pp 77-85. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction*. National Academy of Sciences - National Research Council, Washington, D. C.

Pressey, S. L. 1926. A simple apparatus which gives tests and scores - and teaches. *School and Society*, 23, 373-376.

Rigney, J. W. and Budnoff, I. J. 1962. *The relative efficiency of different combinations of prompting and confirmation for learning a Boolean algebra problem*. Tech. Report No. 34. Prepared for Personnel and Training Branch, Psychological Sciences Division, ONR, Los Angeles by the Dept. of Psychology, Univ. of Southern California.

Scharf, E. S. 1961. *The effect of reinforcement in programmed learning sequences*. Unpublished Master's thesis, Division of Natural Sciences, Univ. of Pittsburgh.

Schramm, W. 1962. *The research on programmed instruction*. An annotated bibliography. Institute for Communication Research, Stanford University.

Sidowsky, J. B., Kopstein, F. F. and Shillestad, I. J. 1961. Prompting and confirmation variables in verbal learning. *Psychol. Report*, 8, 401-406.

Skinner, B. F. 1958. Teaching machines. *Science*, 128, 969-977.

Skinner, B. F. 1959. The programming of verbal knowledge. pp 63-69. *Automatic Teaching: The State of the Art*. E. H. Galanter (ed.), New York: John Wiley & Sons.

Steel, R. G. D. 1961. Some rank sum multiple comparison tests. *Biometrics*, 17, 539-552.

Stevens, A. L. 1950. *Certain special factors involved in the law of effect*. Unpublished Doctoral Dissertation, Department of Psychology, Ohio State University.

Stolurow, L. M. 1961. *A comparative study of methods of programming materials for efficient learning in self-instructional devices*. University of Illinois, Urbana, Illinois.

Stolurow, L. M. and Lippert, H. 1963. *Prompting, confirmation and vanishing in the teaching of a sight vocabulary*. Training Research Laboratory, Univ. of Illinois, Urbana, Illinois.

Appendix A

The order in which the *S*s were run

1. CCC T
2. PCP T
3. PPPPPPPPPPPP T
4. PPP T (PP)
5. PCPCPC T
6. --- T
7. --------- T
8. PPP T (PE)
9. PPPPCPCPCCCC T
10. PCPCPCPCP T
11. CCCCPC T
12. PPPPPPPPP T
13. PPPPPP T
14. CCCCPCPCP T
15. PPPPCP T
16. PPPPCPCPC T
17. CCCCPCPCPPPP T
18. ------------ T
19. PCPCPCPCPCPC T
20. ------ T
21. PCP T
22. PPP T (PP)
23. CCC T
24. --- T
25. PPPPPPPPPPPP T
26. PPP T (PE)
27. PCPCPC T
28. PCPCPCPCP T
29. --------- T
30. PPPPPPPPP T
31. PPPPCPCPCCCC T
32. CCCCPCPCP T
33. CCCCPC T
34. PPPPCPCPC T
35. PPPPPP T
36. ------------ T
37. PPPPCP T
38. ------ T
39. CCCCPCPCPPPP T
40. PCPCPCPCPCPC T
41. PPP T (PP)
42. --- T
43. PCP T
44. PPP T (PE)
45. CCC T
46. PCPCPCPCP T
47. PPPPPPPPPPPP T
48. PPPPPPPPP T
49. PCPCPC T
50. CCCCPCPCP T
51. --------- T
52. PPPPCPCPC T
53. PPPPCPCPCCCC T
54. ------------ T
55. CCCCPC T
56. ------ T
57. PPPPPP T
58. PCPCPCPCPCPC T
59. PPPPCP T
60. CCCCPCPCPPPP T
61. --- T
62. PPP T (PE)
63. PPP T (PP)
64. PCPCPCPCP T
65. PCP T
66. PPPPPPPPP T
67. CCC T
68. CCCCPCPCP T
69. PPPPPPPPPPPP T
70. PPPPCPCPC T
71. PCPCPC T
72. ------------ T
73. --------- T
74. ------ T
75. PPPPCPCPCCCC T
76. PCPCPCPCPCPC T
77. CCCPC T
78. CCCCPCPCPPPP T
79. PPPPPP T
80. PPPPCP T
81. PPP T (PE)
82. PCPCPCPCP T
83. --- T
84. PPPPPPPPP T
85. PPP T (PP)
86. CCCCPCPCP T
87. PCP T
88. PPPPCPCPC T
89. CCC T
90. ------------ T
91. PPPPPPPPPPPP T
92. ------ T
93. PCPCPC T
94. PCPCPCPCPCPC T
95. --------- T
96. CCCCPCPCPPPP T
97. PPPPCPCPCCCC T
98. PPPPCP T
99. CCCCPC T
100. PPPPPP T

Appendix B

Instructions to the Subjects

"The purpose of this study is to see how easy or difficult it is for you to learn a letter code. Please listen to the instructions all the way through and then if it's not clear what you are supposed to do, I'll answer your questions.

First I'll explain the general procedure, then I'll tell you specifically what you're to do.

In this study, each of the first ten letters of the alphabet has a particular code line associated with it. Your job is simply to learn the code line for each letter. Then you'll be tested to see how well you know them.

Each letter in this code is represented by a bent line that connects three dots. For example, this code line represents the letter X. This code line represents the letter Y. Notice that the pattern of dots is always the same, it is only the bent code lines that differ.

There's no system to the code. The code lines were assigned to the letters in a random fashion, so don't waste your time looking for a system. You just have to learn the code line for each letter, one at a time. Remember that the pattern of dots is always the same, it is only the code lines that differ. Only the first ten letters of the alphabet will be used and each letter is represented by a bent line that connects three dots and only three dots. Every letter has a different code line. Don't bother to learn the order in which the letters are presented because the order will always be different.

Now put your name in the upper right hand corner of this booklet. (Pause) If you'll thumb through the sheets, you will notice that they are all the same. Notice also that the boxes on these sheets contain the standard dot pattern and are numbered down the page in two columns: boxes one to five in the first column and boxes six to ten in the second.

We want to find out how much you can learn by actual practice in drawing the code lines. We'll go through all ten code lines several times and you'll practice drawing the code lines on every trial.

In order to determine the effects of practice, you actually have to practice of course. This means that you have to draw a code line for every letter. If you don't know the code line guess at it; but be sure to draw a code line for every letter.

During the training trials, two methods of practicing might be used. It is extremely important that you completely understand these procedures, so listen carefully.

One method that we might use is called a prompting method and it goes like this. Each letter will flash on the screen for a brief interval. The letter will then go off and the correct code line for that letter will then appear for a few seconds. After the code line goes off, you draw that correct code line on the practice sheet. Since this is the first code line in this practice trial, you should draw the code line in box one. You will have only three seconds to draw it, so act quickly. At the end of three seconds, I'll say STOP. When you hear the word STOP, look up at the screen for the next letter. The second letter will appear one second after I say STOP, so look up quickly. The second letter will then appear briefly and it will be followed by its code line. After the code line flashes off, you draw it in box two.

The pattern, then goes like this: letter, code line, draw the code line, stop; letter, code line, draw the code line, stop; and so on. Ar there any questions?

The other method that we might use is called a confirmation method and it goes like this. Each letter will flash on the screen for a brief interval. The letter will then go off and the screen will go blank for three seconds. During this time, you are to draw the correct code line for the letter. Since this is the first code line in this practice trial, draw the code line in box one. After three seconds, I'll say STOP. When you hear the word STOP, look up at the screen and the correct code line will appear. You can now confirm your response. The code line will then flash off and we will have a one-second interval before the next letter appears. Again, the letter will appear briefly, then the screen will go blank. You draw the correct code line in box two in this three-second blank interval. After I say STOP, you look up and the correct code line will appear. The pattern, then, goes like this: letter, draw the code line, STOP, correct code line; letter, draw the code line, STOP, correct code line; and so on. Are there any questions?

As you have probably noticed, the only difference between the prompting method and the confirmation method is the placement of the drawing of the code line. In the prompting method, you see the code line then draw it, in the confirmation method you draw the code line, then you see it."

If the S was in the pre-determined sequence group, the following instructions were then read.

"I will tell you which method you will be using before we start a practice trial, then we'll go through all ten pairs of letters and code lines using that method. Before the next trial begins, I will again tell you which method you will be using. I will also describe the method to be used by saying either 'The next trial will be a prompting procedure and goes letter, then code line, then you draw it' or I'll say 'The next trial will be a confirmation procedure and goes letter, then you draw the code line, then look up for the correct code line.'

Are there any questions before we begin? You will have a practice trials before the final test is given. All right, here we go on the first practice trial which will be a __________ procedure and goes . . . . . ."

If the S was in the SD group, the following instructions were then read.

"I will allow you to choose the method you want to use before we start a practice trial, then we'll go through all ten pairs of letters and code lines using that method. Before the next trial begins, I will again allow you to choose which method you want to use. Before you choose, however, I will again describe your choice by saying 'You may choose either the prompting procedure in which you see the letter, then see the code line, then draw the code line, or you may choose the confirmation procedure in which you see the letter, then draw the code line, then see the correct code line.' Are there any questions?

All right, here we go. You may choose either the prompting procedure in which you see the letter, then see the code line, then draw the code line, or you may choose the confirmation procedure in which you see the letter, then draw the code line, then see the correct code line. Which do you choose?"

If the confirmation procedure was either assigned to or choosen by the S on the first trial, the following instructions were then read.

"On this trial you will be required to draw a code line before you have seen it. Even though you don't know what the correct one is, guess at it. Then look up to see what the correct line should be. Here we go."

After completing the training sequence, the practice sheets were collected and another sheet containing the standard dot pattern was given to the subject. The following instructions were then read.

"That's the end of the training sequence. We will now have a test trial and the procedure will be this. The first letter will appear for a brief interval and the flash off. The screen will go blank and you will have five seconds to draw the correct code line in box one. After five seconds, I'll say STOP, and you look up for the next letter. When it goes off, draw its code line in box two. We will repeat this procedure for all ten letters. Are there any questions? Here we go."

After completing the test, E collected the test sheet and gave S another sheet containing the standard dot pattern. The following instructions were then read.

"On this test sheet, I would like you to draw as many of the correct code lines as you can recall. You may draw them in any order you please, so guess if you can't recall some of them. Are there any questions? You will have five minutes to complete the test."

The final test sheet was collected and the terminal instructions read to the S.

"That's the end of the experiment. Thank you for participating. Before you leave there is one additional item I would like to mention, and it is a very important one. Please do not say anything about this experiment to anyone else. The reason is this. We are planning to continue this research here at State College for several months and it may be that one or more of the people whom you know will be asked to participate. As you can well imagine, if anyone who is a subject knows beforehand what the experiment is about, it would bias the results. I would appreciate it then if you would say nothing about any part of this study to anyone. Okay? Thank you."

- -

Appendix C

Raw data from experiment A

Table 1

C scores for all groups on test one

Level of Training	FE	CE	RS	FP	SD
			Groups		
3	3	1	2	1	3
	3	2	0	2	4
	2	1	0	1	2
	2	2	0	5	1
	3	0	1	4	1
6	6	4	1	9	3
	7	7	7	2	3
	6	2	8	5	7
	10	1	0	4	5
	2	1	2	4	3
9	5	5	8	8	10
	9	7	4	2	10
	4	6	4	0	6
	4	8	6	5	9
	4	4	10	5	3
12	2	3	0	7	8
	4	8	9	6	7
	10	3	7	9	8
	5	7	6	8	6
	2	5	7	4	4

Table 2

D scores for all groups on test one

Level of training	Groups FE	CE	RS	FP	SD
3	3	3	4	3	5
	4	3	4	2	6
	3	4	2	1	5
	4	4	3	6	4
	5	3	2	5	2
6	7	5	4	9	4
	9	9	7	9	4
	8	4	8	6	7
	10	3	3	5	6
	5	2	5	7	4
9	6	5	9	8	10
	9	7	5	3	10
	5	7	6	3	6
	5	9	6	6	9
	5	5	10	5	4
12	2	3	1	7	8
	5	9	9	6	7
	10	5	8	9	8
	5	9	7	8	7
	4	6	8	6	7

Table 3

D scores for all groups on test two

Level of Training	Groups FE	CE	RS	FP	SD
3	3	6	4	6	4
	5	5	4	4	5
	3	5	5	4	5
	4	6	3	7	7
	6	5	3	7	4
6	9	5	5	9	4
	8	8	9	4	6
	9	5	9	5	6
	10	4	3	6	6
	5	4	5	5	8
9	6	6	9	8	10
	8	7	6	3	10
	4	6	7	4	8
	7	9	8	9	10
	6	7	9	6	7
12	4	7	3	7	9
	6	10	9	8	6
	10	6	9	7	7
	7	9	9	10	7
	7	4	8	7	7

Appendix D

Raw data from experiment B

Table 1

C scores for all groups of Experiment B

	Groups			
Level of training	FE	CE	FF	FC
12	9	5	3	10
	6	5	3	8
	2	8	6	10
	7	8	6	7
	8	9	10	4
	4	10	9	8
	2	9	9	8
	10	3	10	5
	10	4	7	7
	0	6	6	8
				7
				5
				9
				6
				4

Table 2

D scores for all groups of Experiment B

	Groups			
Level of training	FE	CE	FF	FC
12	9	6	6	10
	6	6	4	8
	5	8	6	10
	8	8	6	7
	8	9	10	5
	6	10	9	8
	4	9	9	8
	10	6	10	6
	10	5	9	8
	1	6	7	9
				8
				7
				9
				7
				7

The Interacting Effects of Guidance and Anxiety upon Learning

Howard G. Miller, James Brown Grier

and James W. Stamen

In the last decade or so Spence and Taylor have carried on a series of experiments to test the implications of the relationship between performance in learning situations and level of drive. This work and these ideas are summarized by Taylor (1956) and Spence (1958). Drive was identified with high scores on the Manifest Anxiety Scale developed by Taylor from the Minnesota Multiphasic Personality Inventory, (Taylor, 1953). The proposition tested in these experiments was that under a condition of high effective drive (as evidenced by high MAS scores) subjects would perform significantly better on simple non-competitional experimental arrangements involving only a single habit tendency, than subjects under low drive (low MAS scores). The experimental tasks employed in testing this idea have been classical conditioning tasks, chiefly eyelid conditioning. Classical conditioning is represented as a non-competitional situation in which only a single response tendency is acquired (Taylor 1956). A corollary of this hypothesis is that under conditions of high drive (high MAS scores) subjects would perform less well on tasks in which there is competition among a number of responses. In this case the situation will frequently occur in which the correct response is weaker than one or more competing responses, and the incorrect responses will be elicited with some frequency. In addition, under a high drive level new competing responses with relatively weak habit strengths may be brought up to the point of elicitation. Therefore, high drive subjects are likely to perform more poorly than low drive when there are a number of competing responses and where the correct response is not highest and strongest in the response hierarchy. The tasks employed to test this part of the hypothesis have been paired associate learning tasks.

Spence (1958) has summarized these experiments and the related theory and concludes that the experimental evidence from eyelid conditioning and paired associate experiments substantially supports the theory.

Other investigators have formulated other varieties of the general hypothesis relating anxiety to performance. Spielberger (1962) has concluded that performance in college students is related to anxiety particularly in the case of middle ability students, where high MAS scores were related to academic failure.

Other investigators, S. Sarason, Mandler and Craighill (1952), have shown a relationship between test performance and anxiety, with the high anxious persons generally performing less well than low anxious persons. McKeachie, _et al._, (1955) have found evidence that anxiety is related to performance on classroom examinations in college.

In these various tests of the anxiety-performance hypothesis the anxiety level of subjects has been determined in various ways. Spence (1958), Taylor (1956) and Spielberger (1962) measured the anxiety level of subjects by instruments such as the MAS or the Welsh Scales of the MMPI. McKeachie and Sarason induced anxiety by giving ego-related or threat type instructions to subjects. In general as a result of this extensive experimentation it has been well established that anxiety level and performance are related. Although the relationship is a very complex one, there is evidence that high anxiety states are likely to interfere with performance on complex tasks and to improve perfomance on simple tasks.

In this experimentation, simplicity or complexity is thought of as being inherent in the task itself whether that task be eyelid conditioning, paired associate learning, test performance, or academic performance in college. That the simplicity or complexity might inhere in the method of training rather than in the task itself has not been considered. The prompting-confirmation studies such as Cook and Kendler (1956) suggest that simplicity or complexity may also reside in method of training and not only in the task. Under the prompting method in paired-associate learning, the subject is presented the stimulus term, then the response term and then reproduces the response implicitly or explicitly in oral or written form as the case demands. The confirmation condition requires that the subject be presented first with the stimulus term, then make his response to that term and then be presented with the correct response term which confirms the correctness of his response. The parallel between the simplicity and complexity in the research on anxiety and performance and the prompting and confirmation methods of training is clear. Prompting is a method of training in which the subject is presented with a simple learning task; the method reduces to a minimum the amount of response competition during training because the subject is provided with the correct response immediately and other responses are not explicitly elicited. The confirmation condition however permits the subject to respond incorrectly if he happens to come into the experiment with a response hierarchy in which the correct response occupies a weak and low order position.

In general, prompting seems to be superior to confirmation as a method of acquisition. Though several studies have failed to find any statistically significant difference between the two methods, all those studies that have found a significant difference have found prompting to be superior. For a review of the literature see Study - 2 this report. The reasons for the superiority of prompting are not so clear. It is quite possible that some of the superiority might be accounted for by the anxiety-generating character of confirmation as a training method, a phenomenon noted by Cook in his experiments. He reports (informal communication) that there is considerable amount of anxiety observable in subjects apparently generated by the task under the confirmation procedure. This is in contrast to an air of calmness noted in subjects being trained on the same task under the prompting procedure. In the confirmation method the subject is uncertain about the correctness of his response throughout a considerable range of his training procedure, particularly in the early trials. Such uncertainty, as suggested by Cook's observations, may act as a stress situation heightening anxiety and producing some of the effects of high drive on the learning of complex tasks suggested by Spence. Since the confirmation method provides for competition of responses, it may be that the anxiety (drive) generated by the confirmation method contributes to the reduced performance generally noted for confirmation. The contention here is that the anxiety generated by the training method itself may interfere with the effectiveness of the confirmation technique since it would act to increase drive and lead to the elicitation of incorrect responses which at the moment are higher in the response hierarchy than are the correct responses. (Taylor, 1956) Such an effect might be expected for all subjects regardless of chronic anxiety level. However, for those subjects who operate under a chronically high anxiety level (high drive) the effect should be increased due to the interaction of the anxiety level the subject brings to the learning task and the anxiety generated by the training method itself. Therefore we should expect that high anxious subjects should do less well using the confirmation training method than when they are trained under a prompting method; and further that high anxious subjects should be handicapped with reference to low anxious subjects when trained with a confirmation method. Measurement of the degree of emotionality in subjects while being trained under confirmation and prompting procedures should be possible using GSR as the indicator of degree of emotionality or drive. Such measurement should show a lowering of electrical resistance, that is higher drive levels, for subjects being trained with the confirmation procedure. Furthermore it should be possible to demonstrate that highly aroused (anxious) subjects learn better using the prompting procedure.

This report describes a two part experiment carried out to explore the relationship between level of anxiety and the prompting and confirmation methods of training. In the first of these the level of anxiety was determined by the use of the Taylor Manifest Anxiety Scale. The efficacy of prompting and confirmation as methods of training was tested comparing the performance of high anxious subjects and low anxious subjects using both training methods in a paired associate learning task. In the second part of the experiment high anxious subjects were subjected to stress created by ego involving instructions, one half being trained under prompting methods and one half under confirmation methods on a paired associate learning task. During this second study subjects' psycho-physiological behavior was measured during the entire training and testing period by recording their skin resistance on a Stoelting "Deceptograph". Non-specific GSR's were recorded and used to determine the level of emotionality of the subject as developed by Silverman, Cohen and Shmavonian (1959). Thus we were able to determine the waxing and waning of emotionality in a given subject as he was confronted by the various demands of the learning task and its environment as well as to relate emotionality as exhibited by skin resistance to these independent indicators used in these experiments, MAS scores, and stress induced by ego involved instructions.

To summarize, the experiments were carried on to determine the following:

a) The relationship between anxiety level as determined by the Taylor Manifest Anxiety Scale and Prompting and Confirmation training methods in paired associate learning.

b) The effect of stress applied by means of ego involving instructions in the learning performance of high anxiety subjects using prompting and confirmation training methods.

c) The level of emotionality in subjects being trained on a paired associate learning task using prompting and confirmation training methods under conditions of stress.

Subjects.- Subjects were male college students currently enrolled in a course in introductory psychology at North Carolina State of The University of North Carolina at Raleigh. Subjects were assigned to the high anxious group if they were in the top 20% of the distribution on the MAS and to the low anxious group if they were in the bottom 20% of the distribution. All high MAS scores were over 20 and all low MAS scores were under 8.

Procedure.- Subjects were trained on a paired-associate learning task in which the stimulus term was one of the first ten letters of the alphabet and the response term was a code line drawn connecting 3 dots in a non-symmetrical but constant pattern of seven dots (see appendix). Subjects' task was to respond by drawing the correct code line to the stimulus of the letter. There were 10 items in the list. The prompting training method involves projecting the stimulus term on the screen for two seconds. Then the response term, the code line, appears on the screen for two seconds. Following this there is a blank period lasting five seconds in which the subject is to draw the correct code line in the pattern of dots provided. The confirmation procedure requires that the 2-second exposure of the stimulus term be followed by a five-second period during which the subject is to draw the correct code line. This period is followed by the projection of the correct code line so that the subject may determine the correctness of his response.

In the first part of the study, there were four groups of 10 subjects each, two high anxiety groups with one trained under the prompting method, and one by the confirmation method and two low anxiety groups similarly trained.

The training procedure under each of the experimental conditions consisted of 3 training trials and then a test trial until five tests were completed. Consequently in the first part of the experiment there were 15 training trials and five test trials as shown in Figure 1. In the second part of the experiment there were 12 training trials and 4 test trials.

In the second part of the experiment only high anxiety subjects participated. All subjects were put under stress by means of instructions which were ego involving. They were told that performance was highly related to learning ability and that their performance would be carefully watched and compared with other college students around the country. (See appendix) These subjects were put under continued stress during the experiment. They were reminded that their performance was being observed continually and that they were always to do their best. The experimenter was critical of their performance saying, after tests were run, that their efforts were inadequate and that they would have to do better if they expected to make a decent record. These critical remarks were made in an authoritarian and deprecating manner and the general atmosphere present was threatening and unpleasant.

	Prompt	Confirmation
High Anxious	$PPPT_1$ $PPPT_2$ $PPPT_3$ $PPPT_4$ $PPPT_5$	$CCCT_1$ $CCCT_2$ $CCCT_3$ $CCCT_4$ $CCCT_5$
Low Anxious	$PPPT_1$ $PPPT_2$ $PPPT_3$ $PPPT_4$ $PPPT_5$	$CCCT_1$ $CCCT_2$ $CCCT_3$ $CCCT_4$ $CCCT_5$
High Anxious Stress	$PPPT_1$ $PPPT_2$ $PPPT_3$ $PPPT_4$	$CCCT_1$ $CCCT_2$ $CCCT_3$ $CCCT_4$

P = Prompting Trial
C = Confirmation Trial
T = Test

Figure 1. Design for experiment. All groups contained 10 subjects each.

Each subject in the second part of the experiment was connected to a Stoelting No. 22501 Deceptograph polygraph and his galvanic skin response was recorded throughout his performance in the experiment. The Deceptograph was modified to increase the maximum sensitivity so that a 10 ohm resistance change in the subject resulted in a one millimeter pen deflection. The polygraph was attached to the subject with stainless steel electrodes 1.5 x 2 ins. in size, coated lightly with electrode paste and taped securely, one on the sole of each foot. An experienced operator attended the console of the polygraph at all times and kept a record of all events during the experience on the polygraph recording paper.

The GSR response was not the usual base resistance but the number of 1 mm. fluctuations on the record. Such fluctuations are referred to by Silverman, Cohen and Shmavonian as "non specifics." The frequency of these fluctuations is a function of the general arousal level of the organism and therefore reflects the general arousal level of the organism. The technique is described in detail in Silverman, Cohen and Shmavonian (1959). Arousal level is high during the existence of an anxiety state, and anxiety has been cited as having drive properties (Spence, Taylor). The increase in the number of non-specifics may therefore be taken to indicate the degree of drive existing in the subject.

Prior to the beginning of the training and testing sessions subjects remained in a quiet state for 20 to 30 minutes to settle down their psychophysiological activity and to establish a base GSR. The stress-inducing instructions are then given to the subject and the training sessions are begun exactly as was the case in the situation described where no stress or GSR measurement occurred.

Results

Our expectation was that high-anxious subjects would do less well when trained by confirmation methods and better when trained by prompting methods on a paired associate learning task. The reason for this was the simplicity and non-competition of responses which is characteristic of the prompting method and the competition of responses which occurs with the confirmation method. Low-anxious subjects, on the other hand, should give similar test performances under prompting and confirmation methods of training since they,not being under a high drive, will not be affected in their behavior by simplicity or complexity of the learning task.

Figure 2 shows the compared performance between high-anxiety subjects when trained by prompting and confirmation methods. In all tests prompting shows as the superior method of training although significance at or below the .05 level is achieved only for tests 2 and 3, as determined by a Kruskal-Wallis one way analysis of variance, (Siegel, 1956).

Figure 3 shows the comparison between prompting and confirmation for low-anxiety subjects. This comparison shows a general superiority for the prompting method of training on all tests, as was the case for the high-anxiety subjects. The only significant difference in this series of tests, again using the Kruskal-Wallis test, is for tests 1 and 2, although the remaining three tests are in the expected direction. Apparently the difference between prompting and confirmation for these subjects is due to the general superiority of prompting over confirmation rather than because of the effect of anxiety on the performance of these subjects. One additional comparison between prompting and confirmation is possible with the available data; namely, the comparisons between high-anxious subjects under stress where prompting and confirmation methods of training were employed. This comparison is shown in Figure 4. Again the differences were in the expected direction with prompting superior to confirmation although, except for Test 1, the differences between the two methods are not significant.

The effects of anxiety can be further examined in Figures 5 and 6 which show the relationship between high-anxious and low-anxious subjects when trained under both prompting and confirmation methods. Subjects who are prompted show no differences at all whether they are high-anxious subjects or low-anxious subjects or whether they are subjected to experimental stress (figure 5). The curves for these various classes of subjects are remarkably close together on all tests. No statistical significance emerges for any comparisons between groups.

The performance of anxious subjects when trained by confirmation methods presents slightly different and a somewhat ambiguous picture. Low-anxious subjects perform better than high-anxious subjects from Test 2 onward although the differences between the two groups are not significant. Such a difference was expected according to the hypothesis. The performance of the stressed high-anxious group under the confirmation condition was superior to both other groups to the extent that significance was reached when compared with the high anxious group on Test 3. Such differences were clearly not predicted by the hypothesis earlier stated.

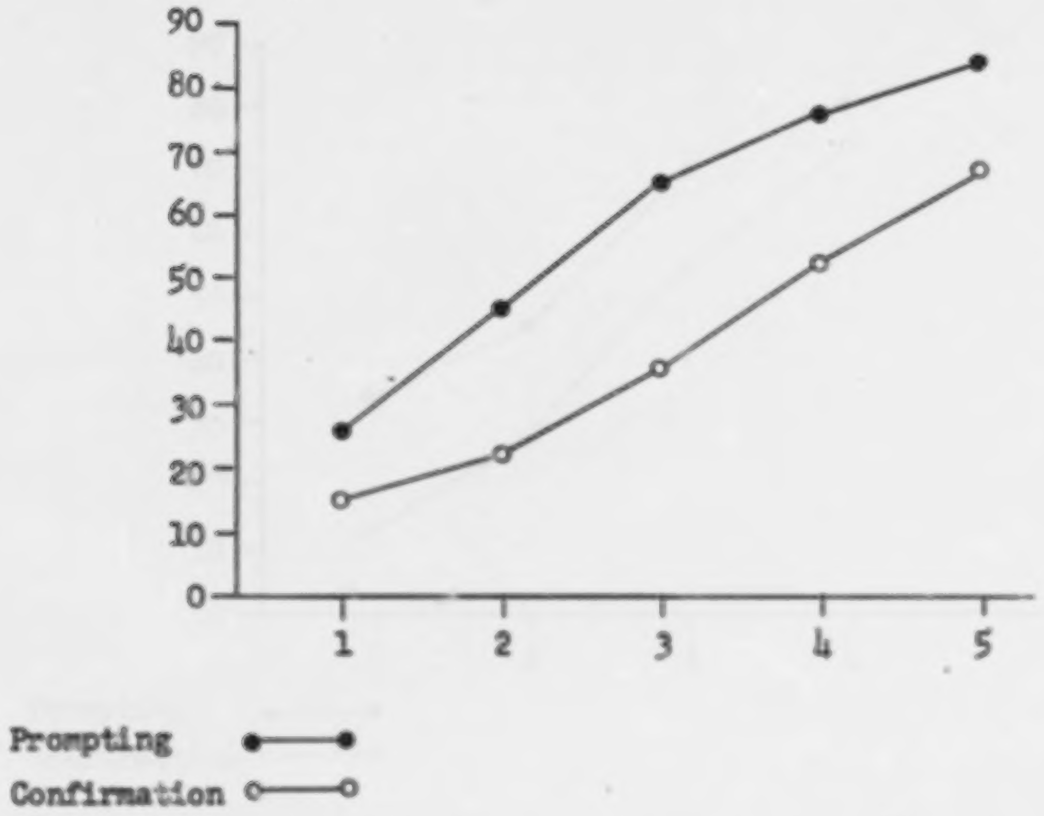


Figure 2. High anxiety groups (no stress) prompting vs. confirmation

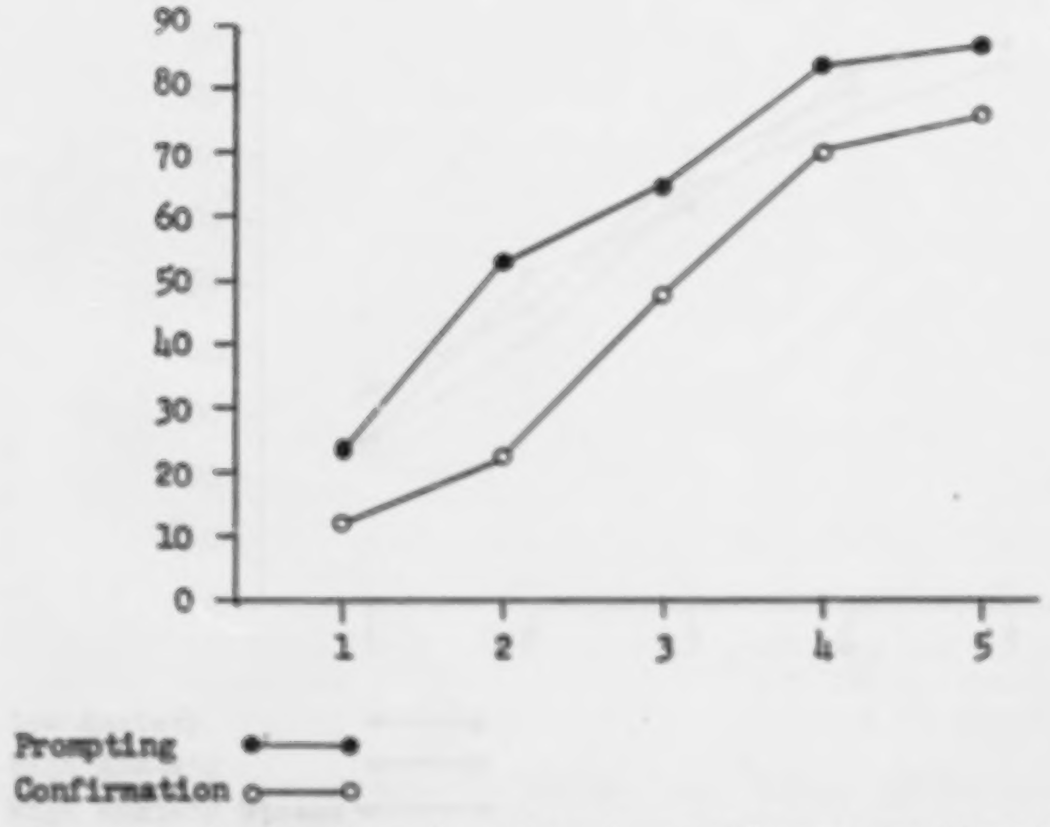


Figure 3. Low anxiety (no stress) prompting vs. confirmation

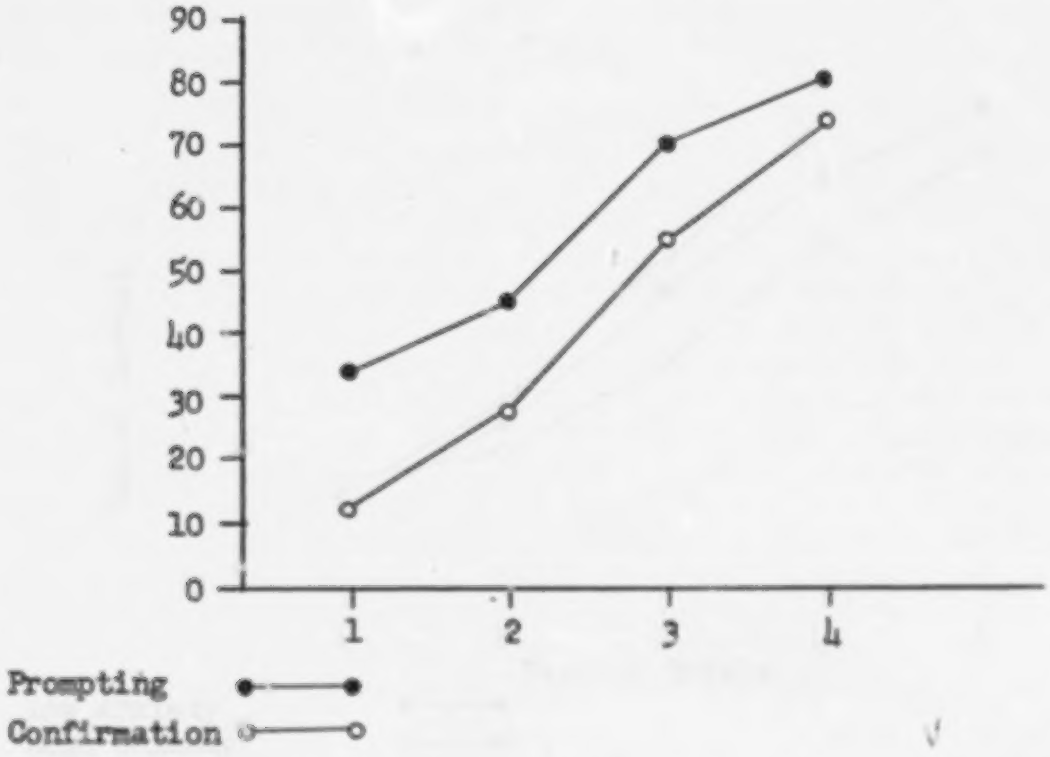


Figure 4. High Anxiety Stress prompting vs. confirmation.

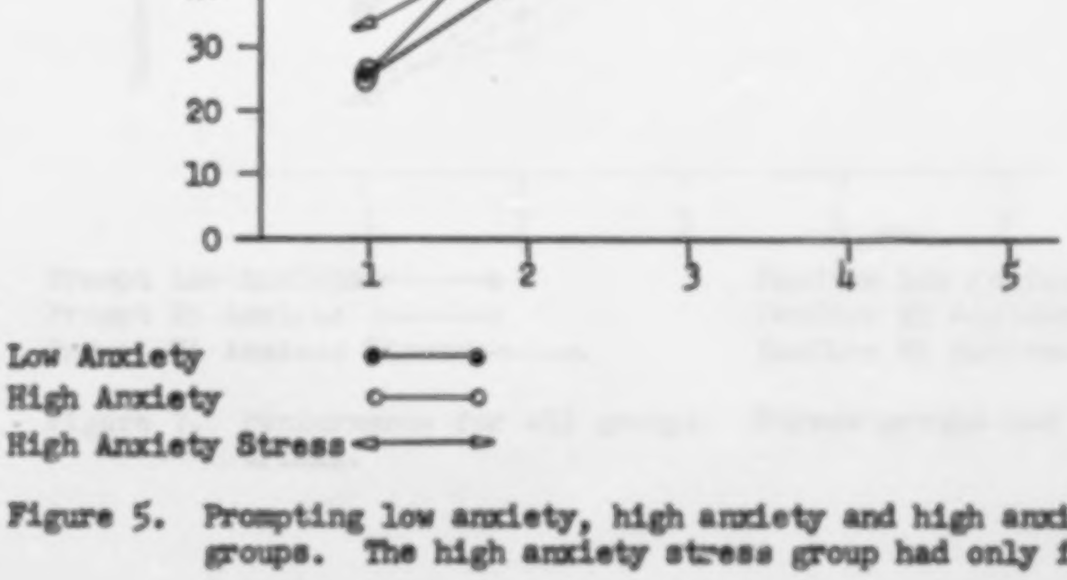


Figure 5. Prompting low anxiety, high anxiety and high anxiety stress groups. The high anxiety stress group had only four testing trials.

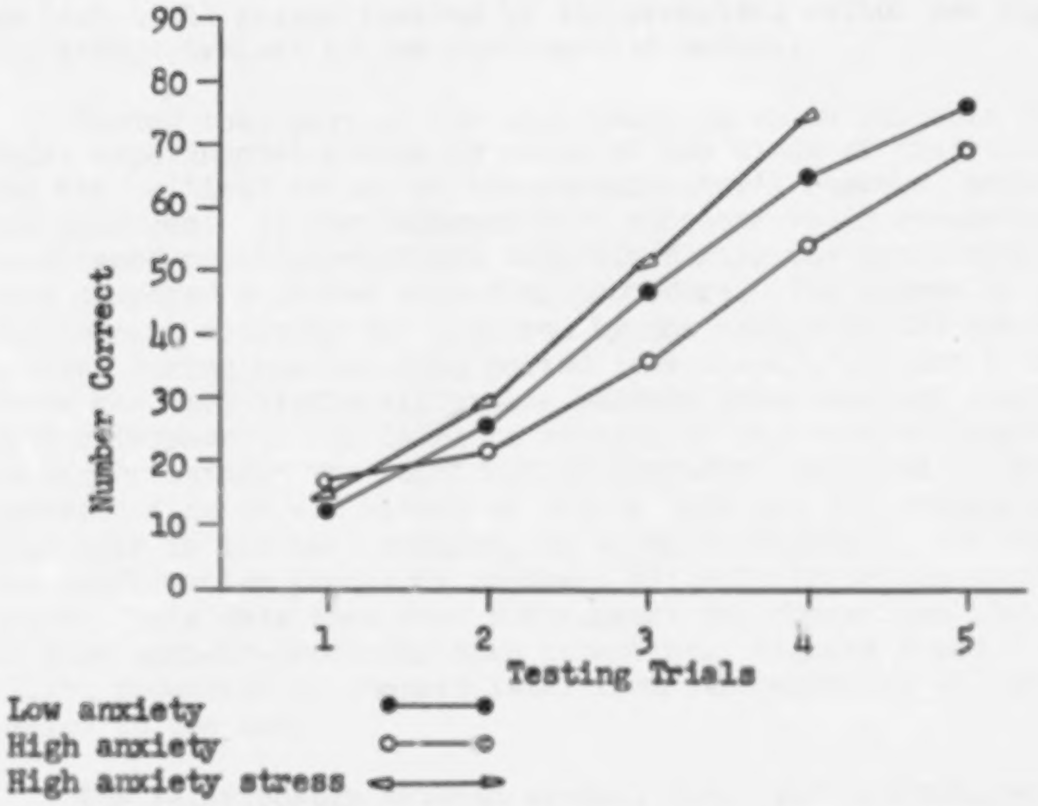


Figure 6. Confirmation, low anxiety, high anxiety and high anxiety stress. The high anxiety stress group had only four testing trials.

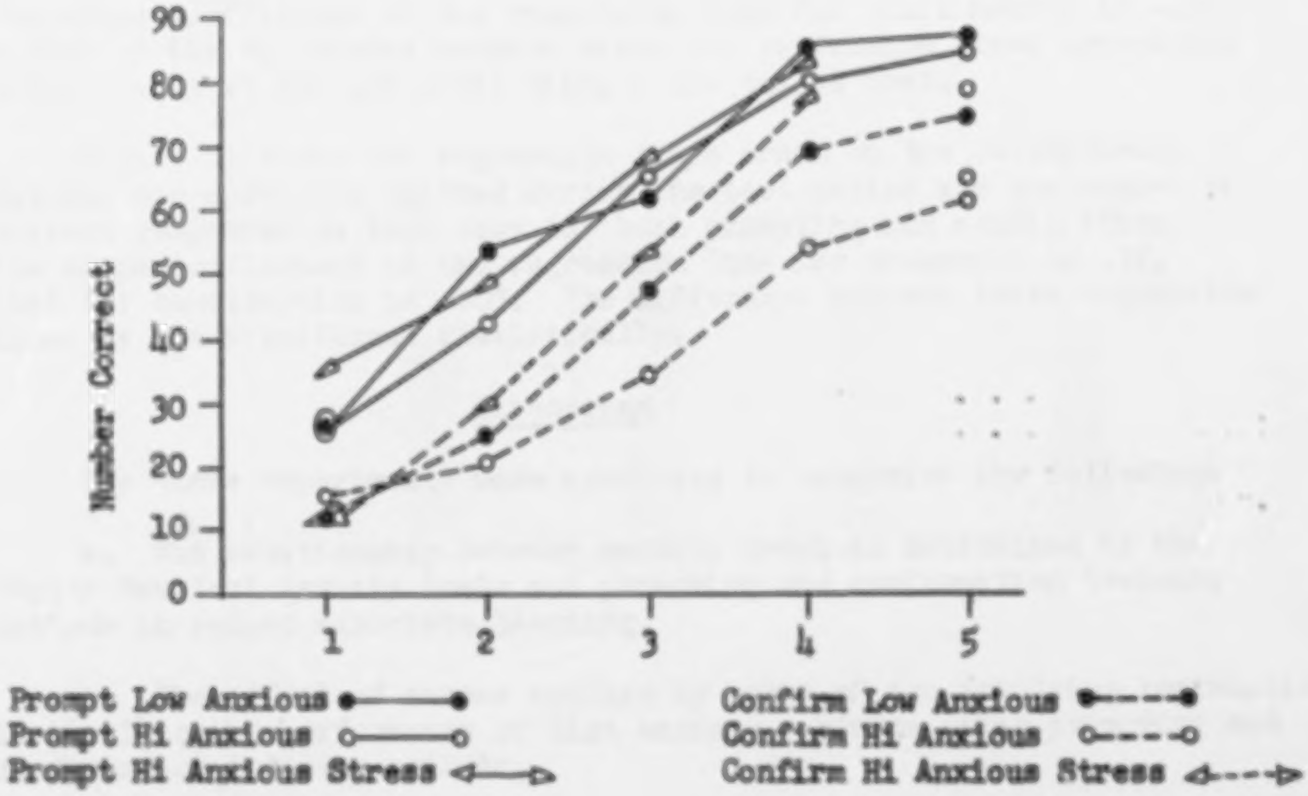


Figure 7. Performance for all groups. Stress groups had only four test trials.

Figure 7 shows all the curves for all six groups for comparative purposes. The notable feature emerging from this comparison is that by Test 4 all groups trained by the prompting method are superior to all groups trained by the confirmation method.

During that part of the experiment in which subjects were placed under experimental stress by means of the kinds of instructions given and the critical nature of the experimenter's remarks, skin resistance was measured. It was expected that subjects would demonstrate heightened psychophysiologic activity during the confirmation procedure when compared with the prompting procedure. The degree of psychophysiologic activity was measured by the number of GSR non-specifics emitted during the learning period (see above). Figure 8 shows that there was very little difference between prompting and confirmation with reference to the level of arousal or psychophysiologic activity during training. The confirmation procedure produced slightly fewer non-specifics in all blocks of trials than did the prompting procedure. Similarly in the test periods, as shown in Figure 8, the subjects under the confirmation procedure produced slightly fewer non-specifics for all tests. This data then does not support the contention that confirmation is more anxiety-producing than prompting. Figures 8 and 9 also show a slight reduction in arousal level from the beginning of the experimental period to the end.

The relationship between arousal level and learning as reflected in the number of correct responses is shown for both confirmation and prompting in Figure 10. The slope coefficient of the regression line drawn on the relationship between number of correct responses on the test immediately following the training period and number of non-specifics during training when the prompting method is used is .21. The slope coefficient of the regression line for confirmation is -.01. A test of the difference between those two regression lines approaches significance at the .05 level using a one tailed test.

Figure 11 shows the regression lines drawn on the relationship between non-specifics emitted during the test period and the number of correct responses on that test for both prompting and confirmation. The slope coeffecient of the regression line for prompting is .32, that for confirmation is -.03. The difference between these regression lines is not significant statistically.

Discussion

The above experiments were conducted to determine the following:

a. The relationship between anxiety level as determined by the Taylor Manifest Anxiety Scale and prompting and confirmation training methods in paired associate learning.

b. The effect of stress applied by means of ego involving instructions in the learning perfomance of high anxiety subjects using prompting and confirmation training methods.

c. The level of emotionality in subjects being trained on a paired associate learning task using prompting and confirmation training methods under conditions of stress.

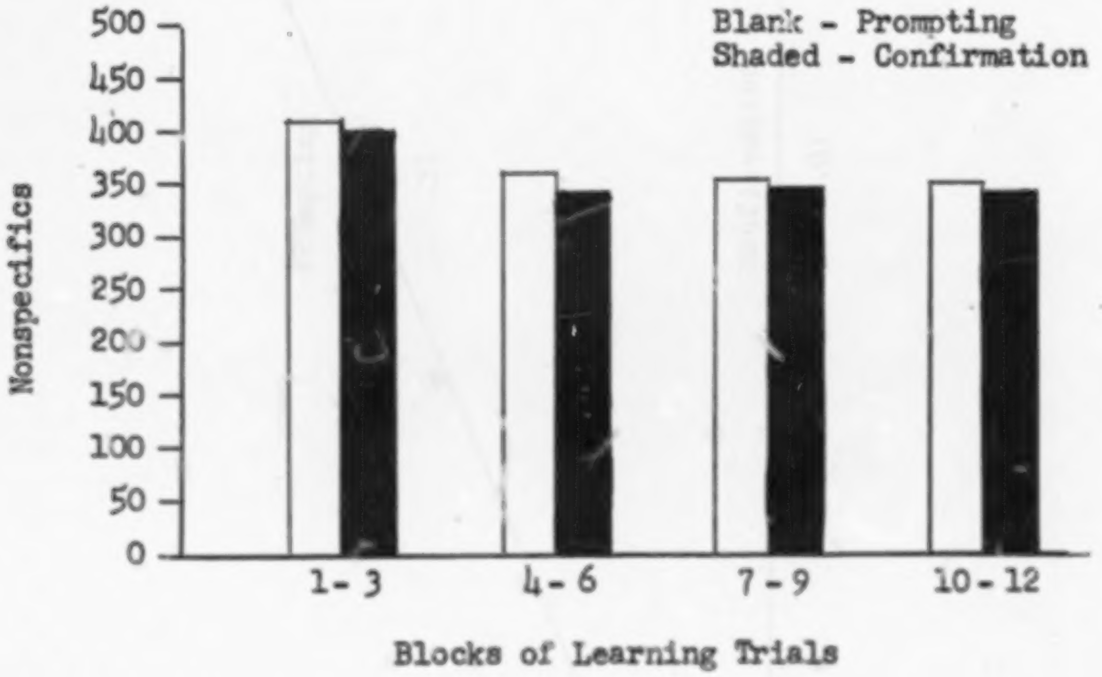


Figure 8. Number of nonspecifics emitted during learning trials

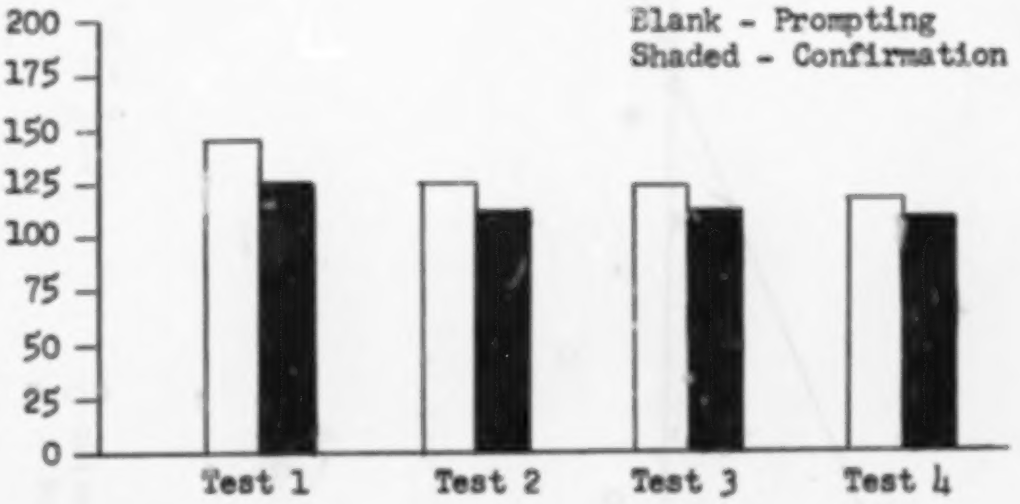


Figure 9. Number of nonspecifics emitted during tests

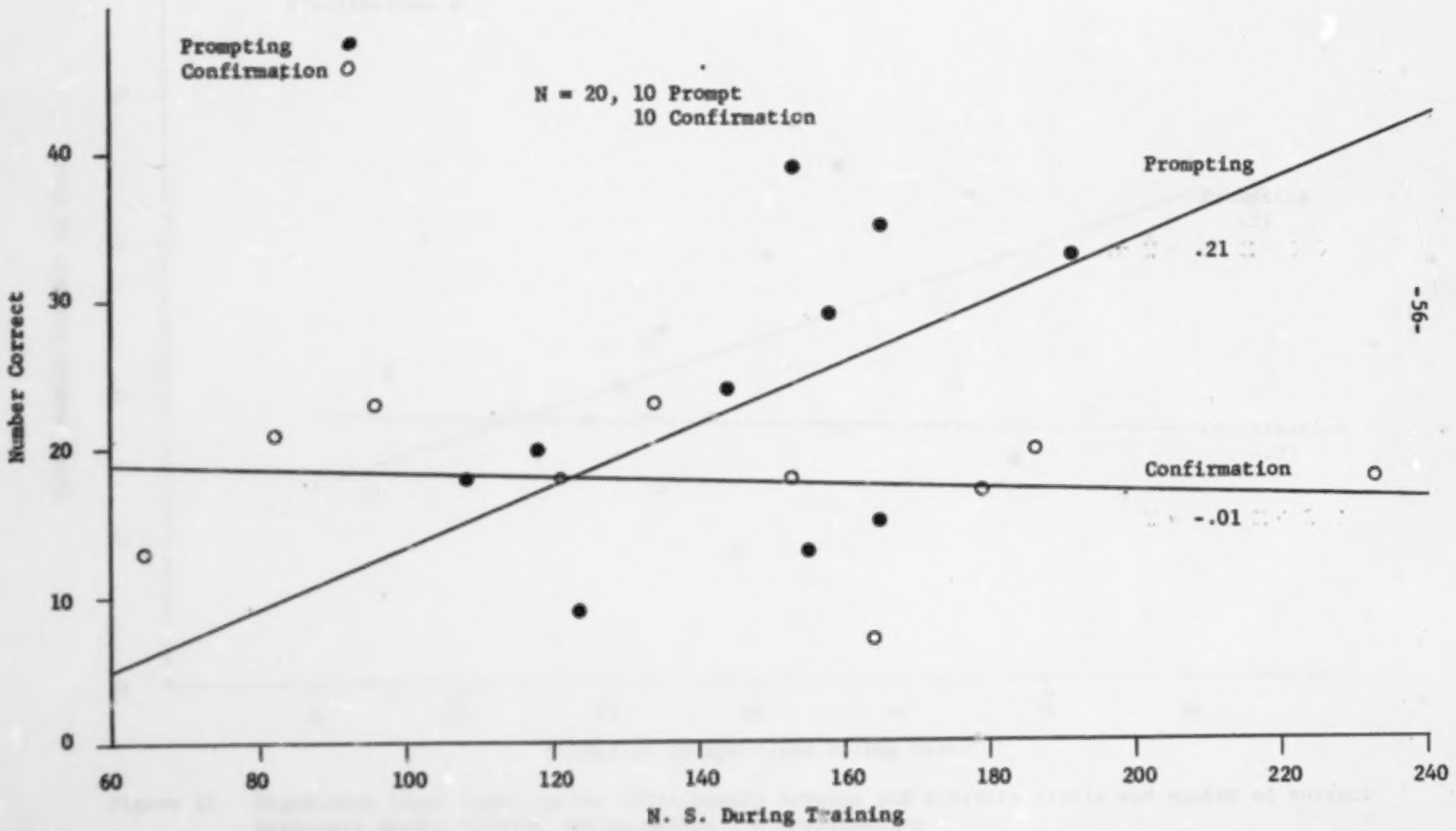


Figure 10. Regression lines drawn on the relationship between GSR activity levels and number of correct responses for prompting and confirmation during training

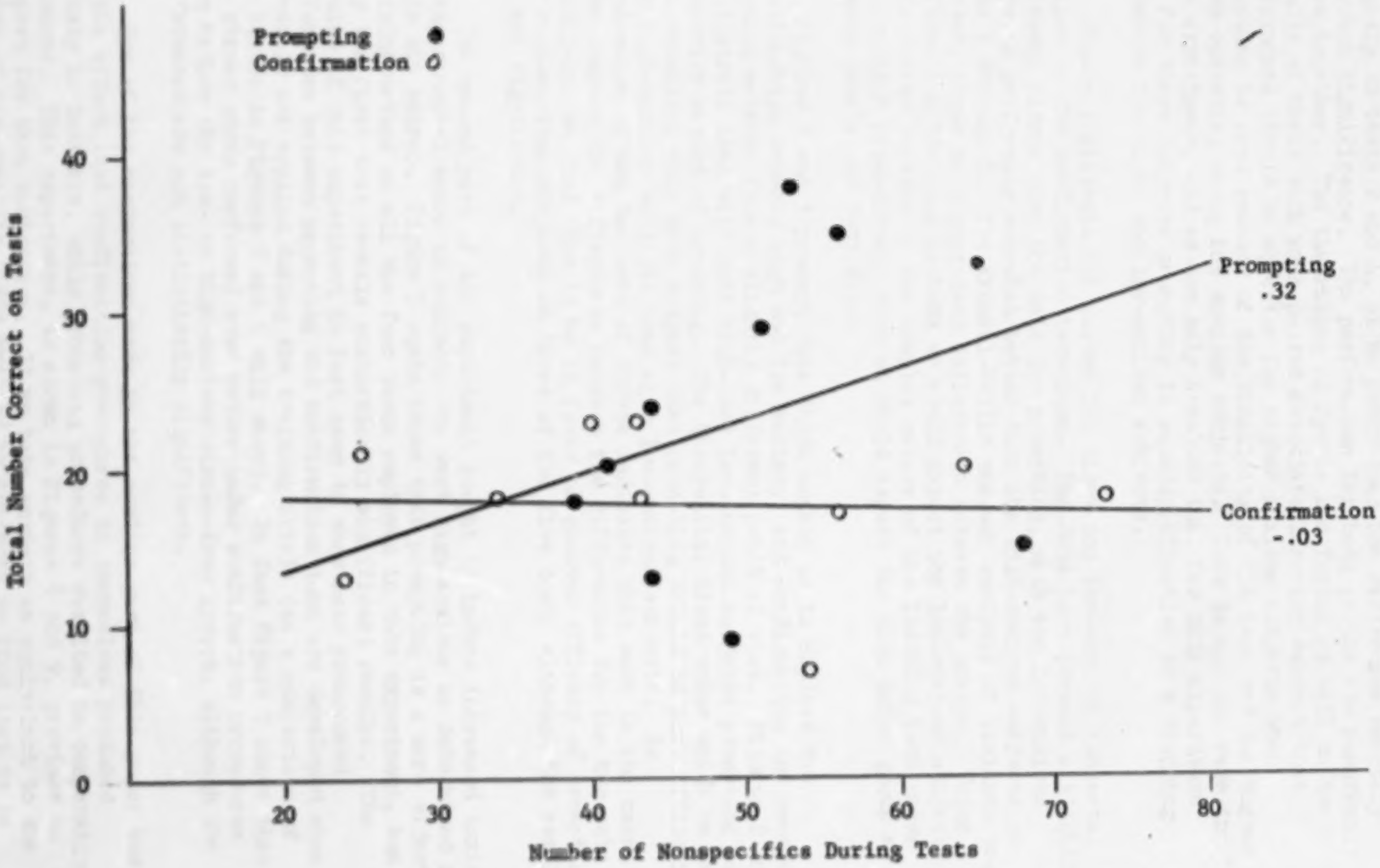


Figure 11. Regression lines drawn on the relationship between GSR activity levels and number of correct responses during testing for prompting and confirmation

Figure 6 shows the curve comparing the effectiveness of prompting with low-anxious and high-anxious subjects. While performance deviates slightly on tests 2 and 4, at no point does the difference remotely approach significance. The performance for both groups are remarkably close together. The theorizing of Spence and Taylor as well as the results of their work with paired associate learning suggest that performance should be superior for higher anxious subjects when prompting is used because of the simplicity of the task and the higher drive operating among the anxious subjects. Such is not the case in this experiment, and we can only conclude that for this experiment and for these subjects prompting is equally effective as a training technique for high- and low-anxious subjects.

Figure 7 presents the curves for high- and low-anxious subjects. trained by the confirmation technique. The data here present a slightly different picture than the data for prompting, with the low-anxious subjects performing womewhat better than the high-anxious subjects on tests 2 through 5. The Kruskal-Wallis one-way analysis of variance test, however, shows no significant differences between the groups. Using confirmation training methods we would expect the low-anxious subjects to do better because of the complex nature of the learning task under confirmation procedures. Here we would expect the high drive group to do more poorly, as they did.

Figures 2 and 3 present data which enable us to explore the relationships between high and low anxiety, and confirmation and prompting training methods from a slightly different point of view. Figures 2 and 3 illustrate that with both high- and low-anxious subjects prompting is a superior method of training. The theoretical ideas under which we have been examining this data suggest that prompting should be more efficacious with high-anxious subjects than with low-anxious subjects. An examination of the two sets of curves indicates that such is the case. If we compare the differences between the differences for the two sets of subjects we find them to be in favor of greater efficacy of prompting for high-anxious subjects on three of the five tests although the results are not significant.

The second part of the experiment sought to induce increased anxiety by experimental means in subjects who were high-anxious as determined by their MAS scores. Figure 5 again shows that prompting is a more effective training method on all the four tests employed in this experiment, but only the first test reveals statistically significant results. The results of this experiment in fact seem to show less pronounced differences between prompting and confirmation than are developed when stress is not applied during the training trials (as a comparison of the curves in Figures 2 and 5 will show). In fact Figure 7 shows that the stress group performed even better under confirmation procedures than either the low- or high-anxious stress-free groups, although the differences are not statistically significant.

One of the contentions made in the earlier part of this paper was to the effect that confirmation procedures in themselves produced anxiety in learners, while prompting procedures resulted in comparative calmness. This experiment, as shown in Figures 8 and 9, provides no support for that contention. If we take anxiety as equivalent to the number of non-specific GSR's emitted by subjects we find that it is slightly less on all trials for learning and testing for confirmation procedures than for prompting procedures.

When all is said and done what can we say about the relationship between anxiety (high drive) and the use of prompting and confirmation training techniques in the learning of a paired associate task?

First, it turns out that prompting is always superior to confirmation regardless of whether the subjects are high-anxious or low-anxious according to the Manifest Anxiety Scale. Second, prompting is superior to confirmation whether or not subjects are learning under a stress condition. Third, confirmation does not seem to generate more anxiety, as measured by GSR non-specifics, than does prompting as a training method. Fourth, there does seem to be a relationship between method of training and degree of anxiety since the regression analysis shows an increasing number of correct responses as psychophysiologic activity increases when prompting is the training method. When confirmation is the training method increased psychophysiologic activity is not related to number of correct responses on the learning task.

Therefore it seems quite clear that we cannot explain the greater effectiveness of prompting over confirmation as a training method on the basis of the anxiety factor since these experiments show far greater differences between the two methods than the presence of anxiety could possibly explain. Yet enough effect of anxiety is demonstrated to suggest at least some relationship between the factors. Explanations as to why differences are not significant, as the theory predicted they would be, are likely to lead to post hoc specious reasoning; a risk which must be run however. The chief reason suggested here is the simplicity of the task. Given the relatively small differences in anxiety level, attenuated by the passage of time between the administration of the MAS and the experimental period, plus the relative simplicity of the task it seems unlikely that the effects of high and low drive had much opportunity to operate effectively in this experiment. We need to experiment with more complex learning tasks which will occupy more time and demand a fuller involvement, intellectual and emotional on the part of the subject. With data from such experiments we should be able to gain more insight into the reasons for differences in prompting and confirmation as training methods and for the relationship of anxiety to such training methods.

References

Cook, J. O. and Kendler, T. S. A theoretical model to explain some paired associate learning data. In G. Finch and F. Cameron (eds.) Symposium on Air Force Human Engineering, Personnel and Training Research, Washington, D. C. Nat. Acad. Sci. Nat. Res. Council Publ. 455, 1956, 90-98.

Mandler, G. and Sarason, S. B. A study of anxiety and learning. J. Abnorm. Soc. Psychol., 1952, 47, 166-173.

McKeachie, W. J., Pollie, D. and Speisman, J. Relieving anxiety in classroom examinations. J. Abnorm. Soc. Psychol., 1955, 50, 93-98.

Sarason, S. B., Mandler, G., and Craighill, P. G. The effect of differential instructions on anxiety and learning. *J. Abnorm. Soc. Psychol.*, 1952, 47, 561-565.

Siegel, Sidney *Nonparametric statistics for the behavioral sciences.* McGraw-Hill, New York, 1956.

Silverman, A. J., Cohen, S. I., and Shmavonian, B. M. Investigation of psychophysiologic relationships with skin resistance measures. *J. Psychosomatic Research*, 1959, 4, 65-88.

Spence, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *Amer. Psychol.*, 1958, 13, 131-141.

Spielberger, C. D. On the relationship between manifest anxiety and intelligence. *J. Consult. Psychol.*, 1958, 22

Spielberger, C. D. The effect of manifest anxiety on academic achievement. *Mental Hyg.*, New York, 1962, 46, 420-426.

Taylor, J. A. A personality scale of manifest anxiety. *J. Abnorm. Soc. Psychol.*, 1953, 48, 285-290.

Taylor, J. A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, 53, 303-320.

Appendix A

Instructions for Subjects Under Stress

The purpose of this study is to see how easy or difficult it is for you to learn a letter code. *We have found that the ability to learn these codes is highly related to intelligence and the ability to do college work. We are interested in finding out more about this relationship and we will keep close records on your performance. Do the best you can and try hard.* (15 sec.) Please listen to the instructions all the way through and then if it's not clear what you are supposed to do, I'll answer your questions.

First, I'll explain the general procedure then I'll tell you specifically what you're to do. Remember to try as hard as you can.

Each of the first 10 letters of the alphabet has its own code line. Your job is simply to learn the code line for each letter. Then you'll be tested to see how well you know them.

Each letter in this code is represented by a bent line that connects three dots. For example - now look up at the screen - this code line represents the letter X. This is Q, notice that the pattern of dots is always the same, only the bent code lines are different. There is no system to the code - the code lines were assigned to the letters in a random fashion. So don't waste your time looking for a system. You just have to learn the code line for each letter one at a time. Remember that the pattern of dots is always the same - it's only the code lines that differ. Only the first 10 letters of the alphabet will be used and each letter is represented by a bent line that connects 3 dots and only 3 dots.

Every letter has a different code line. Don't bother to learn the order in which the letters are presented because the order will always be different.

We want to find out how much you can learn without actually drawing the code lines. We'll go through all 10 code letters several times. After the 3rd time through, we'll have a test run. Three practice trials then a test run, 3 practice trials then a test run and so on until everybody has had a chance to learn the code. In order to determine the effect of the practice trials, you will have to draw a code line for each letter during the test run.

The procedure will be this - the letters will flash on the screen one at a time. Each letter will be followed by its code line like this - the first letter will flash on the screen for 2 sec. and then off and the code line for the first letter will flash on the screen for 2 sec. then off. A second black interval will then appear. The second letter will then flash on the screen then off, then the code line for the second letter will flash on then off. The sequence will be this - letter, code line, black; letter, code line, black; letter, code line, black and so on. Are there any questions? We will record and follow your performance throughout this learning task. At the close of the experiment we will compare your learning speed with that of other similar groups. It has been the experience of psychologists that college students in various parts of the country are able to learn all or almost all of the code lines in the number of trials we are using in this experiment. One of the objectives of this experiment is to see how well students in this part of the country perform on this task as compared with students in other parts of the country. (30 sec.)

All ready, here we go.
P_1 222; 222; 222; etc.
That's the end of the first practice trial. Here we go on the second.
P_2 222; 222; 222; etc.
That's the end of the second practice trial. Here we go on the third.
P_3 222; 222; 222; 222; etc.
That's the end of the third practice trial. Now reach into the test sheet folder - that's the light colored one - and take out the first test sheet. It should say T_1 in the upper left hand corner. If it doesn't please raise your hand. (20 sec.) On the test sheet please print your last name, first name and middle initial in the upper right hand corner. (20 sec.) Notice that the boxes contain the standard dot pattern and are numbered down the sheet in 2 columns. On the test run, the letters will flash on the screen, one letter at a time. Just the letters alone now, not the code lines. When the first letter is flashed on and then off you draw its code line in box one. Draw it quickly. At the end of 5 seconds I'll say STOP. You look up at the screen. A short white interval, then a short black interval will appear. Then the second letter will flash on the screen. When it goes off, draw its code line in box two. Draw it quickly. Then a short white interval, then a short black interval will appear and so on. Don't make any changes once you have drawn the line. The procedure will be: letter, draw, stop, white interval, black interval; letter, draw, stop, white interval, black interval.

Be sure to put a code line in every box - don't omit any items. If you don't know the code line, make the best guess you can but put a code line in every box. This is important. Draw the code line quickly and stop when you hear the word stop, then look up for the next letter. Remember the code line should connect 3 and only 3 dots. Any questions?
All ready - here we go.
T_1 2321, 2321, 2321, etc.
That's the end of the test run. Examiner looks at test sheets. Says, "You'll have to do better than that to meet the standards. Now try hard on the next series." (5 sec.) Put the answer sheet back in the folder, under the blank ones. (15 sec.) Now we will start another series of three practice trials. Remember the procedure. Letter, code lines, black space and so on.

Ready, here we go.
P_1 222; 222; 222; 222; etc.
That's the end of the first practice trial. Here we go on the second.
P_2 222; 222; 222; 222; etc.
That's the end of the second, here we go with the third.
P_3 222; 222; 222; 222; etc.
That's the end of the third practice trial. Now reach into the test sheet folder and take out the second test sheet. It should say T_2 in the upper left hand corner. If it doesn't please raise your hand. (15 sec.) Remember the procedure; letter, draw, stop, white interval, black interval; letter, draw, stop, white interval, black interval and so on.

All ready, here we go. Remember this is a test of your ability.
T_2 2321; 2321; 2321; 2321; etc.

That's the end of the test run. You're falling behind, you'll have to do better. (3 sec.) Put the answer sheet back into the folder under the blank ones. Now we will start another series of three practice trials.

Ready, here we go.
P_1 222; 222; 222; 222; etc.
That's the end of the first practice trial. Here we go on the second.
P_2 222; 222; 222; 222; etc.
That's the end of the second practice trial. Here we go on the third.
P_3 222; 222; 222; 222; etc.
That's the end of the third practice trial. Now reach into the test sheet folder and take out the third test sheet. It should say T_3 in the upper left hand corner. It if doesn't raise your hand. Remember the procedure, letter, draw, stop, white interval, black interval and so on. Do your best.

All ready, here we go.
T_3 2321; 2321; 2321; 2321; etc.
That's the end of the test run. "Come on, you can do better than that." (3 sec.) Put the answer sheet back into the folder under the blank ones. Now we will start another series of practice trials.
Here we go.
P_1 222; 222; 222; 222; etc.
That's the end of the first practice trial. Here we go on the second.
P_2 222; 222; 222; 222; etc.
That's the end of the second practice trial. Here we go on the third.
P_3 222; 222; 222; 222; etc.

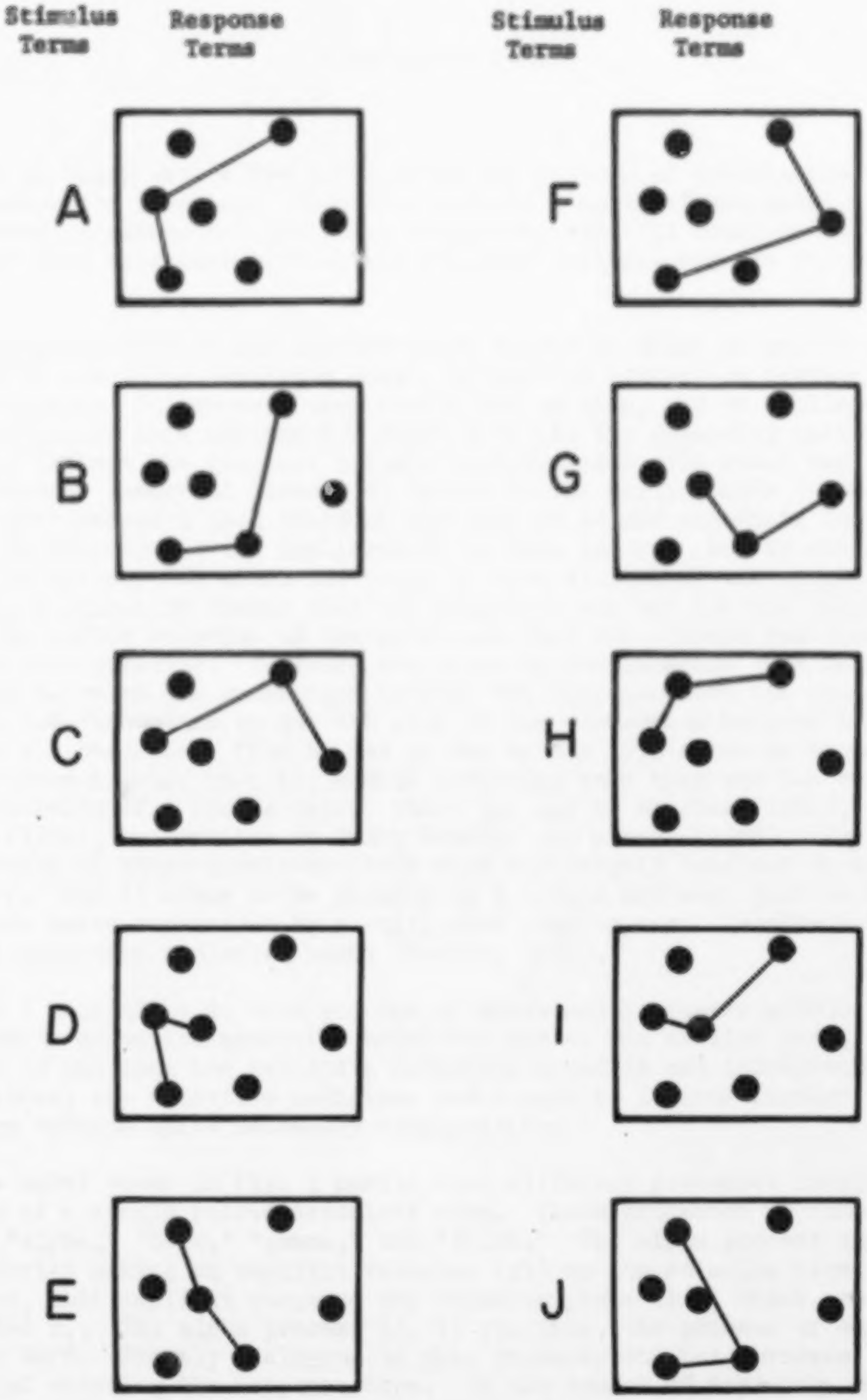
That's the end of the third practice trial Now reach into the test folder and take out the fourth test sheet. It should say T_4 in the upper left hand corner. If it doesn't, raise your hand. Remember the procedure; letter, draw, stop, white interval, black interval and so on.

All ready, here we go.
T_4 2321; 2321; 2321; 2321; etc.
That's the end of the test run. You're still behind. There's only one more test and then a final test. Let's see if you can get up to normal on the next test. (6 sec.) Put the answer sheet back into the folder under the blank ones. Now we will start another series of practice trials. All ready, here we go.
P_1 222; 222; 222; 222; etc.
That's the end of the first practice trial, here we go on the second.
P_2 222; 222; 222; 222; etc.
That's the end of the second practice trial, here we go on the third.
P_3 222; 222; 222; 222; etc.
That's the end of the third practice trial. Now reach into the test folder and take out the fifth test sheet. It should say T_5 in the upper left hand corner. If it doesn't, raise your hand. Remember the procedure; letter, draw, stop, white interval, black interval and so on. Ready, here we go.
T_5 2321; 2321; 2321; 2321; etc.

That's the end of the test run. Put the answer sheet back into the folder under the blank ones. We will now take a short pause before continuing.

We will now take a final test trial. This is your last chance now, do the best you can. (3 sec.) The procedure will be this. The letter will appear for 2 seconds then go off. A four second white interval will then appear. During this time you are to draw the code line on the sixth test sheet. After 4 seconds I will say STOP, look up for the next code line which will appear after a one second black interval. The procedure will be letter, draw, stop, black; letter, draw, stop, black and so on. Are there any questions? Take the test sheet labeled T_T from the test file. Raise your hand if you do not have it.

Ready, here we go.
T_T 241, 241, 241, 241, 241, etc.


Stimulus Terms
Response Terms
Stimulus Terms
Response Terms
A
F
B
G
C
H
D
I
E
J

Theoretical Implications of Response Analysis in Paired Associate Learning

John Oliver Cook

Let me begin with a few words about the history of associative theories of paired associate learning. Then I'll exhibit a multiprocess model, after which I'll present a scheme for analyzing responses, and I'll conclude by trying to point out some relations between the response analysis and the theoretical model.

Throughout most of its history we've tended to think of paired associate learning as something involving only a single S-R connection between a stimulus and a response. Things were very simple for us then, and we exploited the possibilities of this unitary S-R model with all the ingenuity that we could command. Perhaps the greatest triumph achieved with this model was the generalization theory of Eleanor J. Gibson in the early 1940's (Gibson, 1942). For the next decade a good share of the work in paired associate learning was devoted to teasing out the implications of this theory. But by the middle 1950's the unitary S-R model had begun to seem altogether too *simpliste*. The realization gradually dawned that the response term was not the response but simply the second stimulus of the pair, and that the subject had somehow got to learn this stimulus. To those who spoke an S-R language this meant that there had to be an S-R connection between the response term and the response. But this S-R connection is not the same as the connection between the stimulus term and the response. Thus we see in the middle 1950's the development of multi-process models, that is, models involving more than one S-R connection in the learning of a single pair. There was one by Mandler (1954), one by McGuire (1954), and another by Tracy Kendler and myself (1956). Experimental exploitation of these models has been slow and largely confined to their inventors. But it seems to be picking up a little bit now, just as these models are being superseded by a still more complex one. I refer to Slater Newman's selective mediation model (Newman, 1960).

Now I would like to show you one of these multi-process models. This is not Newman's selective mediation model but one of the earlier ones. However, I hasten to add that the selective mediation model is not incompatible with this one. Rather, the selective mediation model adds to it some further complexities, and in my opinion quite necessary complexities.

The model shown in Fig. 1 posits four different processes involved in the learning of a single paired associate item. These processes or connections are labeled "alpha," "beta," "gamma," and "delta." The alpha process consists of the subject's making an implicit response (r1) to the stimulus term. Like all responses, this implicit response has stimulus properties, which are designated s_1. The alpha process is, if you like, the process of encoding the stimulus term. Exactly analogous to this process, the beta process is the process of encoding the response term. In the course of training, either in the experiment or out of it, this second encoding (r_2) becomes attached to the stimulus properties (s_1) of the first encoding. This is the hook-up or gamma process. The delta connection is a decoding process. Once the subject has encoded the response term -- that is, so to speak, once he knows what the correct response is -- he then has to produce it. This is what I mean by saying that the delta process is a decoding process.

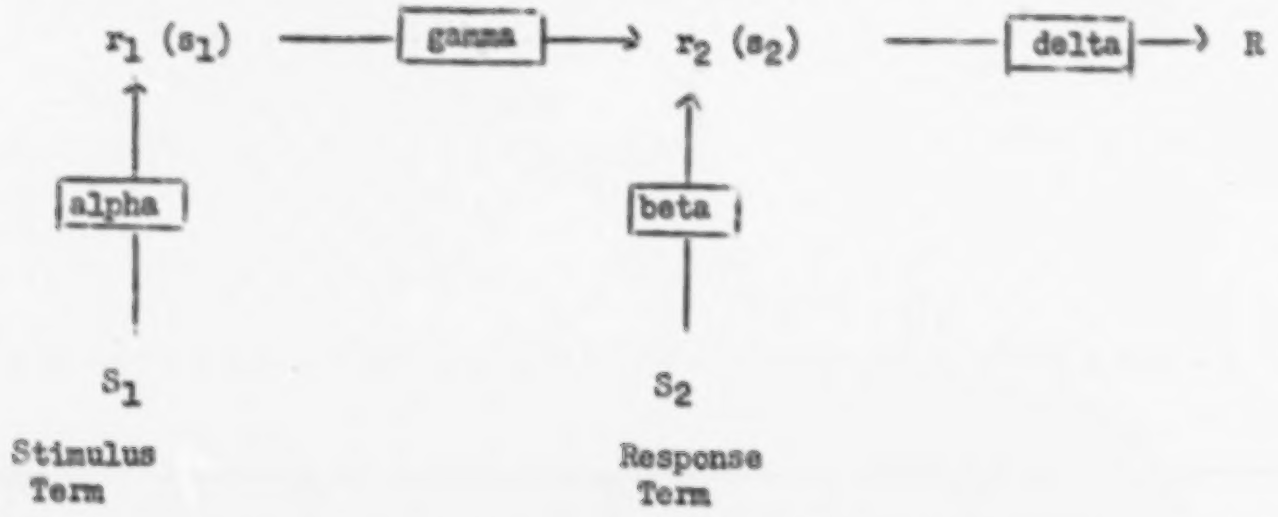


Figure 1. A multi-process model of paired associate learning. (After Cook and Kendler, 1956)

On the test the subject is shown the stimulus term (S_1) and asked to make the response (R). He gets from S_1 to R via these two implicit responses and their stimulus properties namely: r_1 (s_1) and r_2 (s_2).

Now here's where response analysis comes into the picture. If a subject gets a particular item correct, then we assume that all four connections were established. If a subject gets a particular item wrong, however, we know that one or more of these four connections was not established, but we don't know which one or ones. The problem then becomes this: can we develop dependent variables that will reflect the processes pictured in the model? If we can, and to the extent that we can, we will be able to say why this subject did not get the item right. Perhaps he had failed to learn the appropriate response term (that is, failure of the beta connection) or perhaps he had learned the response term but had not learned to hook it up to the stimulus term.. That is, he got it wrong because of a failure of the gamma connection. Perhaps he simply hadn't the skill to produce the response. That is, he had something encoded, but he couldn't decode it in accordance with the scoring rules that were established for that particular experiment.

Now here is a response analysis scheme that may be of some help in dealing with problems of this kind.

T, the total number of response opportunities, can be divided into an L-score (the number of legitimate responses), and X-score (the number of extraneous responses), and an O-score (the number of omissions). The X-score, and perhaps even the O-score, can be further subdivided, but I will neglect them and turn my attention to an analysis of the L-score. To put it crudely, a response is a legitimate one if it happens to be one of the response terms used in the experiment, whether it is evoked by its proper stimulus or by one of the others. As you see in Figure 2, the L-score can be analyzed in two ways. First it can be divided into the D-score, which is the number of different legitimate responses, and the R-score, which is the number of replicates. The second way of analyzing the L-score is to divide it into the C-score, which is the number of correct responses, and the S-score, which is the number of substitution errors.

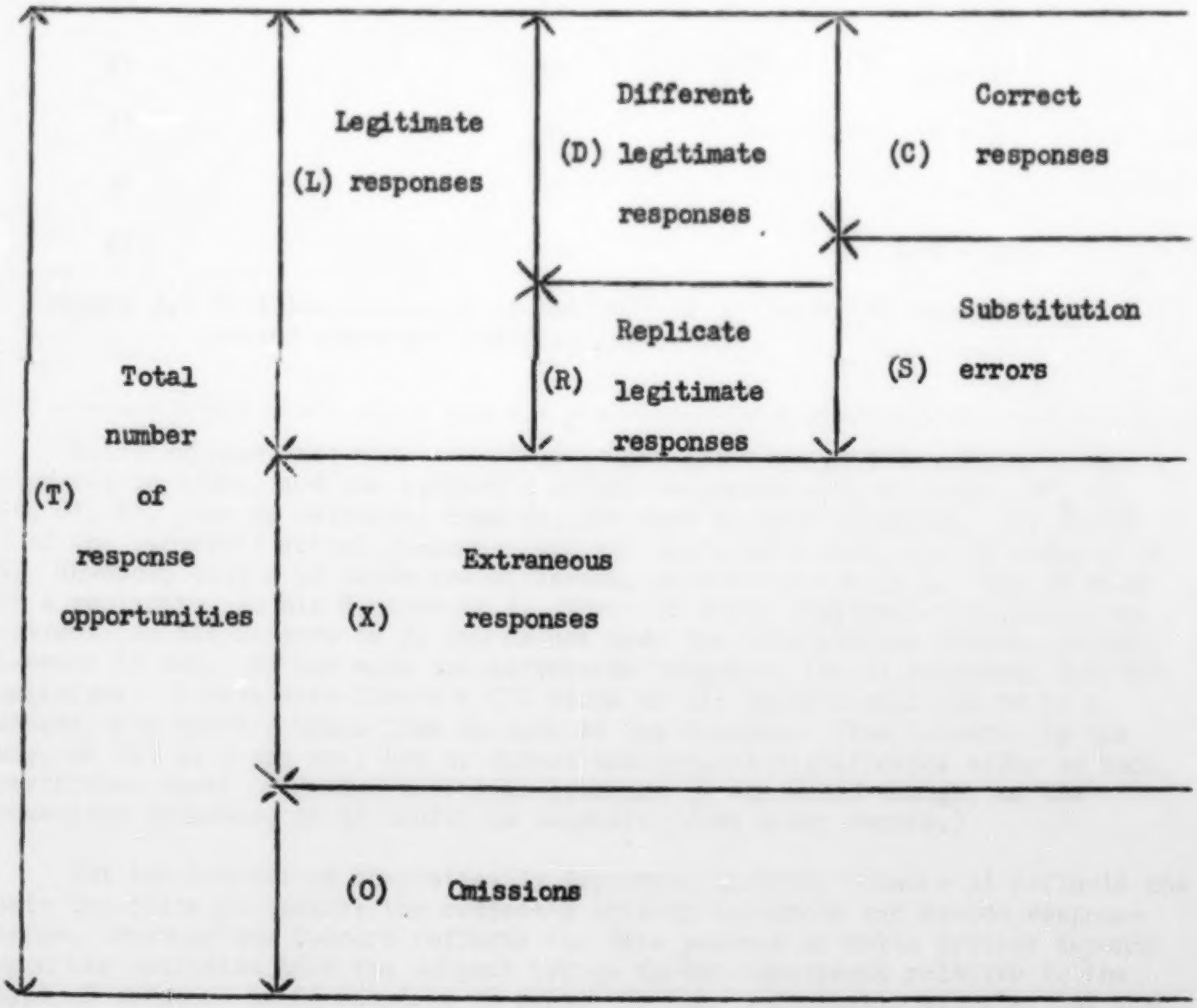

Legitimate
(L) responses
Different
(D) legitimate
responses
Correct
(C) responses
Replicate
(R) legitimate
responses
Substitution
(S) errors
Total
number
(T) of
response
opportunities
Extraneous
(X) responses
(O) Omissions

Figure 2. An analysis of responses in paired associate learning experiments. (After Cook, 1961)

Now here in Figure 3 is an illustration of this response analysis applied to the scores of a single hypothetical subject.

The Correct responses	*S*'s actual responses		*S*'s scores
A'	B'		L = 5
B'	C'		D = 4
C'	C'	L-D =	R = 1
D'	D'	L-S =	C = 3
E'	E'	L-C =	S = 2
F'	--		X = 1
G'	J'		O = 2
H'	--		C/D = .75

Figure 3. An illustration of an analysis of a single *S*'s scores in a paired associate learning experiment.

Let's suppose that the correct responses are the first 8 letters of the alphabet in order, and the subject's actual responses are, in order, B', C', C', D', E', then an omission, then J', and then another omission. The first 5 of the subject's actual responses are all legitimate ones, so his L-score is 5. However, only 4 of these are different, so his D-score is 4. One of them is a replicate, so his R-score is 1. Three of these legitimate responses are correct, so his C-score is 3, and he has made two substitution errors, so his S-score is two. He has made one extraneous response, the J' response, and two omissions. I have also listed a C/D score of .75 which I will get to in a minute, but first I would like to look at the D-score. (The L-score, by the way, as far as I can see, has no direct theoretical significance asfar as this particular model of paired associate learning is concerned though, as the equations indicate, it is useful in computing some other scores.)

But the D-score is theoretically important directly because it reflects the beta and delta processes: the subject's ability to encode and decode response terms. Whether the D-score reflects the beta process or delta process depends upon the abilities that the subject brings to the experiment relative to the type of response he is required to make. To take a fantastic example, but one which I hope makes the point clearly, let us suppose that our subjects are college undergraduates with no skill in drawing. Let us suppose that the stimulus terms are the numbers one through six, the response terms are line-drawing portraits of the first six presidents of the United States, and that the subject must reproduce those drawings. Now the subject may know everything about these materials. He knows the numbers one through six, he can name the first six presidents in order, and he can correctly attach the names to the portraits. The only thing he can't do is draw them. In a situation like this a low D-score can be safely interpreted as a failure of the delta process. The subject can do everything but decode.

But we rarely devise situations like this in the laboratory because we're rarely interested in the motor skill aspect of paired associate learning. However, it is perhaps worth while to point out that it has one, that the model takes account of it in the delta process, and that there is a dependent variable, the D-score, that (given suitable safeguards) reflects the process.

Usually we're at some pains to minimize the motor skill aspect. Even when we use novel terms we try to select materials that we are confident our subjects can reproduce. Under these circumstances if we can rule out the possibility that the D-score is a function of the subject's motor skill, then we can interpret the D-score as a reflection of the beta process. A low D-score, therefore, is an indication of a failure of the encoding process rather than a failure of the decoding process. And a high D-score is an indication that the subject has succeeded in encoding a large number of the response terms.

If the D-score, then, can be interpreted as the number of responses that the subject knows, and if the C-score is the number of responses that the subject knows and can correctly hook up with their proper stimuli, then the C/D score is a measure of the gamma process, or hook-up, because it represents the proportion of responses that the subject knows and can correctly hook up with the proper stimuli. So much for disentangling the beta, gamma and delta processes. I might add that so far I have been unable, and may forever be unable, to disentangle the alpha process from the others without doing R-S testing or something of that kind, which, though possible, is a little inelegant.

In connection with the controversy regarding the theoretical interpretation of substitution errors that enlivened the pages of the Psychological Review a few years ago, (Murdock, 1958; Gibson, 1959; Battig, 1959; Runquist, 1959; Murdock, 1959) it is perhaps worth pointing out that the problem can be attacked in more than one way. The traditional way to explain substitution errors is to attribute them, in a simple and direct fashion, to stimulus generalization: we have many substitution errors if the amount of stimulus generalization is great, and few if it is small.

But, as the equation L-C=S in Fig. 3 indicates, the number of substitution errors is simply the numerical difference between two other scores. It follows logically from the equation that all statements about the magnitude of the S-score are strictly deducible from statements about the magnitudes of the L-score and the C-score. The advantage in establishing this relationship among three variables is that now we don't need three explanations, one for each variable; we need only two, because the explanation for the third variable can be cranked out as a theorem from the other two explanations. We don't need to explain any particular one of these variables; we only need to explain any two. Which two we choose to try to account for is a matter of convenience.

The foregoing idea can be generalized in the utterly obvious statement that whenever we can construct logical or mathematical relations among a set of variables, the number of explanations that we need is always one less than the number of variables.

Let me make three concluding remarks. In the first place, this little excursion has not exhausted the possibilities of this particular scheme of response analysis. For example, it may be that different clinical types (e.g., high anxious and low anxious) though they may make the same number of errors, tend to make different kinds of errors. In the second place, it may be worth pointing out that a response analysis, either one like this or some other kind, may be useful in serial learning or in diagnosing the faults of a teaching machine program. Finally, in the third place, though it may be otiose to mention it, this particular response analysis is not the only possible one nor is it any more legitimate than any other one. And what is true of the response analysis is also true of the theoretical model. It's tempting for the model builder to think that his model is an unposed photograph of Ultimate Reality, a candid camera shot of Nature in the buff. It is with the deepest regret that I must confess it is no such thing. It would be wrong even to call it an approximation. It's simply a conceptual scheme that facilitates the making of a few predictions.

References

Battig, W. F. Comment on "Intralist generalization in paired-associate learning." Psychol. Rev., 1959, 66, 338-339.

Cook, J. O. Response analysis in paired associate learning experiments. In A. A. Lumsdaine (ed.), Student Response in Programmed Instruction. Nat'l. Acad. Sci. - Nat'l. Res. Coun. Publ. No. 943, Washington, D. C., 1961, 367-374.

Cook, J. O. and Kendler, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch and F. Cameron (eds.) Symposium on Air Force human engineering, personnel, and training research. Nat'l. Acad. Sci.- Nat'l. Res. Coun. Publ. 455, Washington, D. C., 1956, 90-98.

Gibson, E. J. A re-examination of generalization. Psychol. Rev., 1959, 66, 340-342.

Gibson, E. J. Intralist generalization as a factor in verbal learning. J. exp. Psychol., 1942, 30, 185-200.

Mandler, G. Response factors in human learning. Psychol. Rev., 1954, 61, 235-244.

McGuire, W. J. A multi-process model for paired associates learning. PhD thesis, Department of Psychology, Yale University, 1954.

Murdock, B. B. Jr. A reply to Battig, Gibson and Runquist. Psychol. Rev., 1959, 66, 345-346.

Murdock, B. B. Jr. Intralist generalization in paired-associate learning. Psychol. Rev., 1958, 60, 306-314.

Newman, S. E. A selective mediation model of paired associate learning. Paper presented at meeting of Psychonomic Society, Chicago, Sept., 1960.

Runquist, W. N. Remarks on "Intralist generalization in paired-associate learning." Psychol. Rev., 1959, 66, 343-344.

Amount of Transfer as a Function of Type of Guidance in Solving Series-Completion Problems

John Oliver Cook and Jerrold Michael Levine

The present experiment had four aims: (1) To test the notion that in problem solving, a more general sort of guidance will facilitate transfer to a new problem to a greater degree than either very specific guidance or no guidance at all. The thinking behind this notion is that if the guidance is specific to the original problem, then S, in facing the new, or transfer problem, will have a great tendency to repeat exactly the same response that solved the original problem; whereas, if the guidance is more general in nature and deals with what the two problems have in common, then the S's performance on the transfer problem should be facilitated. (2) To see whether guidance on one problem diminishes the tendency to persist in attempts to solve a subsequent, difficult - and, in this case, insoluble - problem. The reasoning here is that if the solution of a problem has been made easy through guidance, S may give up quickly on a more difficult problem. (3) To see what effect guidance has on the length of time required to solve a reversal problem; that is, one involving elements that are similar to the original problem but requiring the "opposite" response. A non-reversal shift is one in which the correct response on the subsequent problem is different from the correct response on the original problem, but it is not the opposite response. For instance, to use the Kendlers' example (Kendler and Kendler, 1962), suppose that S is presented with a collection of large and small black cups and large and small white cups. On the original problem, lifting the black cups is the correct response and lifting the white cups is the incorrect response. If, on the subsequent problem, lifting the white cups is correct and lifting the black cups is wrong, then that is a reversal shift. Whereas, if, on the subsequent problem, lifting the large cups is correct, and lifting the small cups is incorrect, then that is a nonreversal shift. Since something at least vaguely analogous to a reversal shift can be constructed in the area of series-completion problems, it may be of some interest to try to find out, what effect guidance on the original problem has on the solution of a subsequent reversal problem. (4) To gather data on the strategies that Ss report employing in solving problems.

Method

Subjects.- The Ss were 60 undergraduate students at North Carolina State of the University of North Carolina at Raleigh. Half of the Ss were male and half female. The Ss were selected randomly and assigned to treatments by a balanced latin square design so that there was an equal number of Ss in each treatment and each treatment followed every other treatment an equal number of times.

Material.- The materials consisted of a stopwatch to record time to solution of the problems and four 3 x 5 blank cards. On each card a problem was typed in black in the center of the card. All problems consisted of six numbers followed by a blank for the seventh number.

The problems were, in order:

Problem 1.	2	3	5	8	13	21	(34)	
Problem 2.	2	3	5	7	10	15	(22)	
Problem 3.	1	3	4	8	11	15	--	(insoluble)
Problem 4.	60	49	11	38	27	9	(18)	

The first problem is simply a Fibonacci series; that is, a series in which each number after the second is the sum of the two previous numbers. The second problem is similar except that alternate numbers are added: i. e., 2 plus 5 equals 7, 3 plus 7 equals 10, 5 plus 10 equals 15, and 7 plus 15 equals 22. The third problem is insoluble. The fourth problem is a reversal of the first problem. Here, instead of a number being the sum of the two previous numbers, it is the difference between the two previous numbers. In the case of each problem S's task was to produce the seventh term in the series.

Experimental Design.- For each of the four problems the design was a non-parametric 2 x 3 factorial analysis of variance as described by Drewes and Bishir (1963). The dependent variable in each case was the time to solution for the problem.

Treatments.- The specific guidance group (SG) was instructed that each number in the series was the sum of two previous numbers. The general guidance group (GG) was instructed as follows: "Each number in the series is a function of some of the previous numbers." The no-guidance group (NG) was given no instructions on how to solve the problem.

Procedure.- An information card was filled out for each S as soon as he entered the experimental room. General instructions describing the experiment and the S's task were then read. Additional instructions appropriate to the assigned treatment were given. Questions were answered by restating the instructions. Four problems were presented, one after the other. If S arrived at a wrong solution he was told to try again and timed until he had correctly solved the problem. There was a ten-minute time limit on problems 1, 2 and 4. Prior to the presentation of problem 3, the Ss were read the following instructions; "The next problem may be too difficult for you. If it is, just say so and I'll give you an alternate problem." There was a 15-minute time limit on this problem. Ss were timed until they gave up or until the time limit was reached. This problem (as stated previously) had no correct solution. Following the correct solution of the fourth problem, Ss were interviewed in order to find out how they went about solving problems 1, 2, and 4. Standard dismissal instructions were then read and the Ss were excused.

Results

The time to solution of each problem was recorded for each S. Since several Ss failed to solve one or more of the problems in the allotted time, non-parametric statistical tests were used to analyze the data. Table 1 shows the median time to solution in seconds for each of the four problems for the six groups. The figures for problem #3 indicate persistence time for the insoluble problem.

The performances of the 30 males and 30 females were compared for each problem under each guidance condition. Mann-Whitney U-tests indicated that there were no significant differences between males and females on any problems under any of the treatments.

Table 1

Median Time to Solution (in Seconds)

	MALES			FEMALES		
Treatments	SG	GG	NG	SG	GG	NG
Problem						
1	12.5	211.0	61.5	15.0	59.0	112.5
2	63.5	74.5	59.5	69.0	249.0	40.0
3	196.5	311.0	195.0	159.5	197.0	176.5
4	30.5	36.5	53.5	35.5	98.0	50.5

Spearman rank-correlation coefficients were computed between problems under all guidance conditions. The correlation coefficients were high enough to indicate that the four problems were not independent. Thus, instead of one over-all statistical treatment, a separate non-parametric analysis of variance had to be computed for each problem. Table 2 shows the analysis of variance for each of the problems.

Table 2

Non-parametric Analysis of Variance for each Problem

	Source	d.f.	H	P
Problem 1	Sex	1	.09	n.s.
	Guidance	2	23.25	< .01
	S x G	2	.80	n.s.
	Total	5	24.14	< .01
Problem 2	Sex	1	-	n.s.
	Guidance	2	-	n.s.
	S x G	2	-	n.s.
	Total	5	3.44	n.s.
Problem 3	Sex	1	-	n.s.
	Guidance	2	-	n.s.
	S x G	2	-	n.s.
	Total	5	2.20	n.s.
Problem 4	Sex	1	1.10	n.s.
	Guidance	2	8.56	< .02
	S x G	2	3.61	n.s.
	Total	5	13.27	< .01

For each problem, sex as a main effect was not significant, confirming the results of the U-test. Guidance effects were significant in problems 1 and 4 only, and there were no significant interactions for any problem. Mann-Whitney U-tests indicated that in problem #1 specific guidance (SG) was significantly better than the other two treatments. This finding is, of course, not surprising. In problem #4 specific guidance (SG) was significantly better than general guidance (GG) and no different from the no-guidance (NG) treatment. Comparisons between the Female - GG cell and each of the other cells, as shown in Table 3, suggest that the significance of the guidance variable in problem #4 is attributable to the poor performance of the F-GG cell.

Table 3

Mann-Whitney U-Test Comparisons Between the Female-GG Cell and the Other Cells in Problem #4

F - GG	M-SG	p	.05
F - GG	M-GG	p	.05
F - GG	M-NG	p	.05
F - GG	F-SG	p	.05
F - GG	F-NG	p	.05

Table 4 summarizes the data collected concerning the strategy reported used by each S in solving problems 1, 2, and 4.

Discussion

The results of this experiment are, in general, negative. With regard to the first aim, the results offer no significant support for the hypothesis that a more general sort of guidance on the first problem facilitates the solution of a subsequent problem to a greater degree than either specific guidance or no guidance at all. As shown in Table 2, the analysis of variance for Problem 2 is completely lacking in statistical significance. The trend for Problem 2 in Table 1 is in the direction of indicating that guidance exerts a deleterious effect upon the solution time of a subsequent problem. However, it should be pointed out that these results do not provide any statistically significant evidence that guidance is harmful.

Similarly, with regard to the second objective of the experiment, it can not be said on the basis of these findings that guidance either significantly enhances or significantly diminishes the tendency to persist in attempts to solve an insoluble problem.

As far as the third objective is concerned, the data for problem #4 indicate that guidance did not facilitate the solution of a reversal problem, and in the case of one cell (F-GG), it significantly increased the time to solution. However, it should be pointed out that the present study is not a pure test of the effect of guidance on reversal shifts, for,

Table 4

Strategies that S̲s Reported Using

		Problem 1			Problem 2			Problem 4		
Group		Initial	Second	Final	Initial	Second	Final	Initial	Second	Final
SG-M	1	5	1		5	1	3	4		
	2	5			5			4		
	3	2	1		2	1	3	8		
	4	1			1	3		4		
	5	1			1	3		8		
	6	1			1	2	3	4		
	7	1			1	3		4		
	8	1			1	2	2	4		
	9	1			1	5		4		
	10	1			1	5	3	8		
SG-F	11	1			1	2	3	2	4	
	12	7	2	1	1	5	9	4		
	13	1			1	2	2	2	4	
	14	1			1	2	0	2	8	
	15	1			1	3		8		
	16	1			1	5	3	2	8	
	17	1			1	9		4		
	18	1			1	3		4		
	19	1			1	2	5	4		
	20	1			1	2	3	8		
GG-M	1	2	5	1	2	5	0	2		
	2	2	1		1	3		8		
	3	5	1		5	1	3	8	4	
	4	2	1		1	2	7	2	4	
	5	2	7	1	1	3		8		
	6	1			1	3		4		
	7	7	1		1	7	3	8	4	
	8	2	2	2	2	7	0	4		
	9	2	5	1	1	3		8	4	
	10	5	7	1	5	1	7	8	4	
GG-F	11	1			3			8		
	12	2	2	0	2	9		2	4	
	13	2	5		2	6	0	2	4	
	14	5	2	1	2	9		8		
	15	2	7	1	1	3		8		
	16	1			1	5	7	8		
	17	2	7	0	2	7	0	2	4	
	18	2	1		2	1	3	2	8	
	19	1			1	3		7	4	
	20	2	7	1	1	2	3	2	4	

Table 4 (Continued)

Strategies that Ss Reported Using

		Problem 1			Problem 2			Problem 4		
Group		Initial	Second	Final	Initial	Second	Final	Initial	Second	Final
NG-M	1	5	1		5	1	5	8		
	2	2	7	0	2	7	9	7	4	
	3	2	5		2	5		2	4	
	4	2	5		3			4		
	5	1			1	3		4		
	6	2	5		2	5	9	8	4	
	7	1			1	3		8		
	8	2	5		2	5		8		
	9	2	1		3			2	4	
	10	2	1		2			4		
NG-F	11	5	7	1	5	1	3	5	4	
	12	2	1		2	1	3	8		
	13	7	6	1	1	6	3	5	8	
	14	5	2	0	2	9		2	8	
	15	2	5	1	2	1	3	8		
	16	5	7	1	2	1	3	8		
	17	2	1		3			8		
	18	2	9		2	9		2	8	
	19	2	1		1	2	3	8		
	20	2	5	1	2	3		6	4	

Code

1 - Added successive numbers

2 - Took differences and looked for pattern

3 - Added alternate numbers

4 - Subtracted successive numbers

5 - General addition

6 - General subtraction

7 - Geometric

8 - Number 11 was clue

9 - Illogical or lucky guess

0 - Gave up or failed

while problem #4 is a reversal shift with respect to problem #1, it is a non-reversal shift with respect to problem #2. As an additional impurity we have the intervention of problem #3, the insoluble problem.

With respect to the fourth aim of the experiment, that of gathering data on the strategies reported employed by Ss in solving problems, one or two perhaps interesting, though statistically non-significant, patterns emerge from an examination of Table 3. However, the theoretical interpretation of these patterns remains obscure.

The first pattern, shown in Table 5, concerns the effect of the guidance instructions (treatments) administered before problem #1 upon the choice of an initial strategy on problem #2.

Table 5

Strategy 1 (Adding Successive Numbers) in Problems #1 and #2

	TREATMENTS SG	GG	NG
No. of Ss reporting using strategy 1 in solving problem #1	19	16	13
No. of Ss reporting using strategy 1 as initial strategy on problem #2	17	10	4
Percentage	89.5	62.5	30.8

In choosing strategy 1 as the initial strategy in problem #2, if nothing were operating but the effect of having successfully used strategy 1 in solving problem #1, then the percentages of Ss shown in Table 5 should all be equal, differing only by chance. But, instead, although the percentages do not differ significantly, they form a regular, declining curve, and the differences are large. This finding suggests that guidance instructions on one problem have an effect upon the selection of an initial strategy for a second problem that is over and above the affect attributable to the successful use of that strategy in solving the first problem.

Table 6 shows that neither the instructions nor solving problem #1 by strategy 1 had any effect upon the number of Ss reporting using strategy 3 in solving problem #2.

Table 6

Strategy 1 (Adding Successive Numbers) in Problem #1, and
Strategy 3 (Adding Alternate Numbers) in Problem #2

	TREATMENTS		
	SG	GG	NG
No. of Ss reporting using strategy 1 in solving problem #1	19	16	13
No. of Ss reporting using strategy 3 in solving problem #2	12	11	12
Percentage	63.2	68.8	92.3

The numbers are virtually identical: 12, 11, and 12. The percentages exhibit a statistically non-significant, but quite regular, trend. Apparently general guidance did not make strategy 3 more available to the Ss in the GG treatment; hence, it is understandable that they did not solve problem #2 faster than the other Ss.

Table 7 bears on the question of whether guidance in solving problem #1 makes strategy 4 more readily available in solving problem #4, that is, upon the reversal-shift effect.

Table 7

The Effect of Solving Problem #1 by Strategy 1 (Adding Successive Numbers) Upon the Choice of Strategy 4 (Subtracting Successive Numbers) as the Initial Strategy in Problem #4

	TREATMENTS		
	SG	GG	NG
No. of Ss reporting using strategy 1 in solving problem #1	19	16	13
No. of Ss reporting using strategy 4 as initial strategy on problem #4	9	1	2
Percentage	42.1	6.3	15.4

The superiority of specific guidance on problem #1 in making strategy 4 more readily available on problem #4 is large, but not statistically significant.

The obvious and regular, trend in Table 8, if it were statistically significant, would strongly suggest that guidance on problem #1 has an effect upon the ultimate choice of the successful strategy (strategy 4) in problem #4.

Table 8

Strategy 1 (Adding Successive Numbers) in Problem #1, and Strategy 4 (Subtracting Successive Numbers) in Problem #4

	TREATMENTS		
	SG	GG	NG
No. of Ss reporting using strategy 1 in solving problem #1	19	16	13
No. of Ss reporting using strategy 4 in solving problem #4	12	8	5
Percentage	63.2	50.0	38.5

This, indeed, may be the case, but the present data do not convincingly demonstrate it.

References

Drewes, D. W. and Bishir, J. A non-parametric analysis of variance for higher order designs. Manuscript, North Carolina State of the University of North Carolina at Raleigh, Raleigh, N. C., 1963.

Kendler, H. H. and Kendler, T. S. Vertical and horizontal processes in problem solving. Psychol. Rev., 69, 1-16, 1962.

A Generalized "Plan" for Serial Learning

John Oliver Cook, Howard G. Miller,
James Brown Grier, and James W. Staman

"A plan is any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed. (Miller, Galanter, and Pribram, 1960, p. 16).

It is not unknown for experimenters to approach a problem with a particular practical goal in mind and then to find that the results lead to theoretical conclusions. The present research is a case in point. We began with the aim of trying to improve performance in serial learning by the administration of guidance, and we ended, as others have done (Asch, Hay, and Diamond, 1960, and Miller, Galanter, and Pribram, 1960) by calling into question the adequacy of the classical theory of serial learning. We propose in its stead an interpretation of how human beings accomplish serial learning which may be considered an addendum to the classical theory, a modification of it, or perhaps, a different theory entirely.

Our practical concern has been to experiment with guiding the learning process. By "guidance" we mean any condition or arrangement of conditions designed to facilitate mastery of a task. If some condition accidentally happened to prove facilitative we would not speak of it as "guidance" since the condition was not designed to be facilitative. According to the usage proposed here, we can speak of guidance as good or bad, effective or ineffective, etc. That is to say, the effectiveness of guidance is a matter for empirical test; it is not a matter of definition.

Research on guidance in learning has had a curious history. There seems to have been considerable interest in it during the 1920's and 1930's. We have come across, respectively, 13 and 32 articles on the topic published in those two decades. Then interest dropped off sharply; we have found only 6 articles on guidance published in the 1940's. Interest began to revive again in the 1950's (15 articles), probably in response to the practical demands created by the development of simulators, training films, and teaching machines (e.g., Kimble and Wulff, 1961, Lumsdaine, Sulzer, and Kopstein, 1961, and Wulff and Stolurow, 1961).

The reason for the drop during the 1940's is interesting to speculate upon. It was certainly not because guidance had been demonstrated to be an utter failure. Indeed, Carr's (1930) review of the early studies reports several notable successes in the guidance of human learning. However, attempts to apply guidance to rats engaged in maze-learning were characteristically failures. Alonzo (1926), for example, tried, without much success, to improve their performance by tugging them through the maze on a little leash.

The decrease in studies on guidance may have been due in part to the fact that the rat became a favorite experimental subject in the late 1930's for testing the implications of theories, and no one had discovered how to make rats respond well to guidance. But perhaps even more important than thi is the fact that the dominant theoretical formulation of serial learning existing at the time really had no place for guidance: look at Hull's (1930) theoretical model shown in Fig. 1.

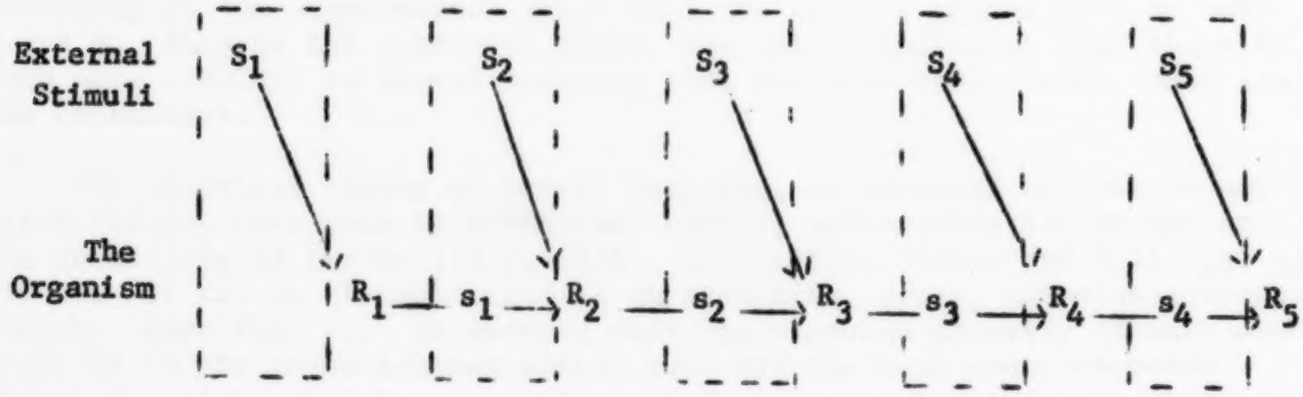


Figure 1. The integration of serial learning. After Hull (1930).

Today it seems as quaint, naive, and simpliste as Aristotle's cosmology. It is the picture of a primitive robot alone with the stimuli. This theory, with its strong emphasis on rote learning, is inhospitable to the notion of guidance.

In serial learning the task is to learn to perform a set of responses in a particular sequence. Thus, as Horowitz (1961) suggests, serial learning consists of two components: (a) S must learn the responses composing the set and, (b) he must learn to perform them in a particular sequence. The responses composing the series may be all of the same kind (e.g., saying the words in a serial list, pressing the buttons of a punchboard maze),or they may be quite different from each other, as in the case of the take-off procedure for an airline pilot. In addition, there is another distinction between different kinds of serial learning tasks. In some, like the pilot's take-off procedure, S has little or no information about whether he has, or has not, omitted a correct response. Indeed, the main problem facing S in such a task may be to insure that every response that should be included is included. In other serial learning tasks, like learning a finger maze, if S is to reach the goal it is virtually impossible for him to omit a correc response. Hence, insuring the inclusion of every correct response is a matte of no concern.

In the present research we are dealing with a task (punchboard maze learning) of the latter type, in which there is relatively little likelihood that S will fail to include a correct response. Moreover, in our task, the responses that make up the series are all of the same sort and are responses that S already knows how to make. We are interested only in how S learns to sequence these responses.

It is perhaps worth mentioning that the responses in serial learning are not the stimuli. Indeed, they are not, in most cases, even duplicates of the stimuli. Consider the nonsense syllable CEV typed in capital letters on the tape of a memory drum. If S is making written responses, and he prints "c e v" in lower case letters, or even uses script letters, his response is considered correct, though it is not a duplicate of the syllable typed on the tape. If S is making vocal responses, his response will be the production of some sound waves. In either case, the response is not a duplicate of the stimulus, but rather a symbolic equivalent of the stimulus according to some equivalence rules that are tacitly agreed upon by both S and E. This is not a trivial point, because it indicates that there is much more involved in serial learning than the classical theory would lead one to believe.

The classical theory of serial learning, as reviewed by McGeoch and Irion (1952), goes back to Ebbinghaus (1885), and reaches its apogee in the theorizing of Lepley (1932, 1934), Hull (1930, 1935), and Hull, et. al. (1940). It can be characterized as an atomistic, overt, stimulus-response theory. (See Fig. 1). It asserts that the stimulus property of each overt response in the to-be-learned series cues off the next overt response. Though it allows for the possibility of remote associations (the third item may cue off the fifth item instead of the fourth one), and perhaps even backward associations (the fifth item may evoke the third one), the basic building block of the theory is still a single, simple connection between an overt response and the stimulus property of another overt response. There is no hint in the classical theory that anything in the nature of a strategy might be involved in serial learning, though the idea that in learning a maze rats develop "hypotheses" was proposed rather early by Krechevsky (1930). Later, Tolman (1948) elaborated the conception that maze learning involved the acquisition of a "cognitive map." Neither Krechevsky's notion nor Tolman's was cast in the form of a stimulus - response theory, and Tolman's view, especially, came under heavy attack from the Hullian classicists.

A "strategy" is essentially a design (good or bad, efficient or inefficient) for achieving a goal. As we will explain later, we conceive of a strategy as an implicit response that cues off a particular set of implicit or overt responses rather than some other set of responses. Perhaps the closest, and, we believe, the earliest approach in S-R terms to the idea of a strategy in serial learning is McGeoch's (1942, p. 94) "mediate association hypothesis," which speculates that " ... spatially remote items in a list might be associated through the connection of each with a third or mediating item. This mediating factor might be anything whatever in the environmental and experimental context, or it might be some common factor in the subject's symbolic repertoire, ..."

Though, as previously stated, there is little place for the notion of guidance in the classical theory, this theory, in conjunction with some findings in other areas of learning, does entail a prediction about the effectiveness of guidance. The reasoning goes like this. As shown in Fig. 1 originally each item in the sequence, operating as an external stimulus, evokes an overt response, but eventually each overt response except the first one becomes conditioned upon the stimulus property of the immediately preceding overt response. When these latter connections have been established, it is only necessary to present to S the first item in the list, and the whole sequence of responses will roll out.

The connection between the stimulus property and the succeeding overt response is an S-R connection, and thus has all the characteristics of S-R connections in general. One of these is that as we lengthen the temporal interval between stimulus and response beyond the optimum interval of about .5 seconds, we make it more difficult to establish the S-R connection.

Evidence on this point comes from conditioning studies (Kimble, 1961, p. 155). Another characteristic is that as we fill up that temporal interval with activities, we make it still more difficult to establish the connection. We have evidence on this point from studies of paired associate learning in human beings (Cook and Spitzer, 1960).

In learning to sequence a set of responses as, for example, in learning a punchboard maze, S has to hunt for each correct response. Inevitably, this lengthens the S-R interval and fills that interval up with activity. A guidance technique, therefore, which shortens the S-R interval and reduces the amount of intervening activity should result in faster learning.

Hawker (1960) achieved some promising results in this direction. He used an electric punchboard maze consiting of ten rows of buttons, with six buttons in each row. One button in each row was arbitrarily selected by E as correct. The pattern of correct buttons was the same for all Ss, and all Ss practiced the maze from bottom to top. Above each button was a small ligh S used a probe which was wired to the maze in such a way that touching a correct button with the probe caused one of the lights to go on. It remained on as long as the probe was in contact with the button.

Hawker employed two training procedures, called "confirmation" and "prompting." In the confirmation treatment S successively touched the probe to the buttons in the first row until he happened to touch the correct button whereupon the light above that button went on, confirming the correctness of his response. The S did the same thing in the second row, and so on through the maze.

In the prompting treatment S first touched the probe to a "start" button centered below the first row of the maze. This caused the light above the correct button in the first row to go on and remain on as long as the contact was maintained. S then touched the correct button in the first row, which caused the light to go on over the correct button in the second row. When he touched that button it lit the light over the correct button in the third row. In this way the lights guided S through the maze, thus shortening the time between successive correct responses (and hence shortening the duration of the S-R intervals) and also eliminating interfering activity during this interval.

The results, as measured by number of trials to criterion, showed that the difference between the treatments was small and statistically insignificant. This was rather surprising, since it had been expected that the prompting treatment would be superior. However, the possibility remained that the failure of the prompting procedure to prove superior might be due to distraction. During the prompting training while the probe was on, say, the fifth button in the second row, the light was on over, say, the fourth button in the third row. The presence of two cues (the probe and the light) on the board at the same time might have produced some distraction that impaired the efficacy of the prompting training.

In the present experiment we wanted to remove this source of possible distraction. To this end, the equipment used was designed in such a way that for example, the light in the second row would not go on until the contact in the first row had been made and broken.

The first aim of the present experiment was to compare this new prompting technique with several other guidance procedures and with the confirmation procedure.

The second aim was to interview our Ss to see if we could get some further insights into how they were going about learning this serial task. Hawker had interviewed his Ss and had learned that they used various strategies. Some used a geometric strategy: they tried to memorize the pathway as a shape. Others assigned numbers to the buttons and tried to memorize the series of numbers. We did not make much of this at the time, but, subsequent to our own experiment the idea of a strategy has come to assume a major role in our thinking about serial learning.

Method

Subjects.- The Ss were 60 male undergraduates in the introductory psychology course at North Carolina State College run during the spring semester, 1962. Ss were selected on the basis of availability and were assigned randomly 10 per treatment, to the first five treatments as they appeared at the laboratory. Ss in the sixth treatment (confirmation) were run later, since approximately two weeks' equipment modification was necessary before this group could be run. Each S was run individually.

Equipment.- The apparatus was an electronic punchboard maze, the face of which contained a 12 x 12 array of pushbuttons, with a jewel light above each button. The path through the maze consisted of one correct button in each row. The correct buttons, which were selected by means of a random number table, were, in order, from the first (bottom) row to the twelfth (top) row (numbering from the left side): 11, 5, 7, 11, 9, 7, 12, 1, 4, 8, 1, 3. Only the lights above these buttons could be activated. The circuitry of the maze was designed in such a way that the maze could be programmed to provide either immediate feedback of the correctness or incorrectness of an overt response (confirmation) or any one of the five types of guidance that constituted the other experimental treatments.

Design.- The design was five-group, one-way analysis of variance for the five guidance treatments. After the addition of the no-guidance (confirmation) group, a six-group, one-way analysis was performed. The dependent variable was the number of trials to reach a criterion of the first of two consecutive perfect runs through the maze.

Treatments.- The Self-Paced Prompting (SelPP) condition was similar to Hawker's prompting condition. The light over the correct button in the first row was turned on by E. S then pushed the button below that light. Upon releasing the button, the light over the correct button in the first row went off, and the light over the correct button in the second row went on, and so on through the maze.

The Confirmation (C) treatment was the same as in Hawker's study. Each S pushed buttons in the first row until he hit the correct one, whereupon the light over that button came on and stayed on until S released the button. Then S did the same thing in the second row, third row, and so on. One trip through the maze was considered one training trial.

In the Simultaneous Display Prompting (SimDP) treatment the lights over all 12 correct buttons came on and stayed on for 36 seconds. This constitute one training trial. (Thirty-six seconds was the mean time required by a pilot group of Self-Paced Prompting Ss for one trip through the maze.) S made no overt button-pushing response during the training trials.

In the Sequential Display Prompting (SeqDP) treatment S made no overt response; he simply observed the lights. The light over the correct button in the first row automatically came on for three seconds and off; then the correct light in the second row came on automatically for three seconds and off; and so on through the maze. The time required for one training trial was thus 36 seconds. S made no overt button-pushing responses during training.

The Sequential Display Prompting with Memory (SeqDPM) treatment was similar to the SeqDP treatment except in two respects: (a) once a light came on, it stayed on until the last (twelfth) light went off, and (b) Ss responded overtly by pushing buttons.

The Automatically-Paced Prompting (APP) condition was just like the SeqDP condition except that during the three seconds that each light was on, S practiced overtly by pushing the button under that light.

Procedure.- After entering the experimental room, S filled out an information card, and then E read the instructions to S while S sat facing the maze. Questions were answered by restating the instructions. S was tested with no feedback after each training trial (one trip through the maze or an equivalent exposure to the lights). The test consisted of S's pointing to the correct lights in the proper sequence while E noted S's choices. S used his preferred hand in both training and test. After reaching criterion S was interviewed to get his account of how he went about learning the maze.

Results

Quantitative.- In terms of number of trials to the first of two consecutive perfect runs, the six treatment means fell into three groups, shown in Table 1.

Table 1

Mean trials to criterion

A		B			C
Sel PP	Sim DP	Seq DPM	C	Seq DP	APP
5.6	5.9	7.3	7.3	8.1	9.9

Using the least significant difference (Steel and Torrie, 1960) there were differences significant at the .01 level between the worst treatment in Group A and the best treatment in Group B, and between the worst treatment in Group B and the one treatment in Group C. There were no significant differences within these groups.

Interview.- As Table 2 shows, three of the 60 Ss in the experiment could not be interviewed because of a conflicting appointment.

Table 2

Interview Data

	N				%
Total N				60	
Not interviewed			3		
Interviewed			57		
Uninterpretable		4			
Interpretable		53			100.
Numerical strategy	49				92.45
Mixed strategy	3				5.66
Geometric strategy	1				1.89

The first 4 Ss interviewed could not give any interpretable clues as to the methods they used in learning the maze, probably because we were not asking the right questions. Each of the remaining 53 Ss was told (if he did not volunteer the information), "Now give me the list of numbers that you used." Forty-nine of these 53 Ss rattled off the list with a speed that makes it difficult to believe that they had learned the maze by some other method and were now translating the path into a list of numbers. We believe that these 49 Ss used a numerical strategy in learning the maze. Three Ss used a mixed, numerical-spatial strategy; that is, they assigned numbers to some of the buttons and coded the locations of the other buttons in terms of their spatial relations to other objects (e.g., Ss left shoulder, a screw on the side of the maze, etc.) One Ss learned a kind of geometric strategy. He "just learned" the location of one or two of the correct buttons, and coded the locations of the others in terms of their directions from the ones he had "just learned;" e.g., "Number 5 ran diagonally [from Number 4]."

Discussion

Implications for classical theory.- The present study began with the practical idea that guidance could improve performance in a serial learning task by reducing the number of errors, which decreases the length of the temporal interval between successive correct overt responses and eliminates almost all of the intervening activity during this interval.[1] This prediction is in accordance with the classical theory of serial learning, and since the results confirmed it, this would ordinarily be construed as supporting the classical theory. However, there are two aspects of this study that suggest that the classical theory gives us a good deal less than a complete account of how serial learning is accomplished.

In the first place, the classical theory does not predict which guidance procedure will be best. More particularly it fails to explain why the Sequential Display Prompting condition is significantly superior to the Automatically Paced Prompting condition. In both treatments the correct lights came on in sequence for three seconds each. The only difference was that in the APP condition *S* responded overtly. Indeed, on the basis of the classical theory it would seem that we would have to predict that the APP treatment would be better. The fact that it is not suggests that the classical theory needs to be modified in the direction of "liberalized" S-R theory as envisioned by such theorists as Miller (1959) and Osgood (1953) so as to include implicit responses as well as overt ones.

In the second place, the classical theory fails to explain the interview data. Should we regard the accounts that *S*s give us about how they learned the maze as simply irrelevant stories that they invent in order to make their behavior appear rational, or should we view the accounts as data that an adequate theory should be able to handle? In both alternatives we have to take the *S*s' accounts seriously because what the *S*s say is behavioral data. The only real alternatives are whether we should devise a separate theory to explain why the *S*s tell us preposterous stories or whether we should try to devise a theory of serial learning that will be compatible with the stories. We opt for the second alternative,and what this second alternative suggests is that *S*s not only make implicit responses, they employ strategies.

For us, the explanatory task is that of constructing a theoretical model of serial learning that includes both implicit responses and strategies Such a model as the one shown in Fig. 2 may be regarded as an S-R rendition of a cognitive map. It also conforms to the definition of a "Plan," as formulated by Miller, Galanter, and Pribram (1960), with one small cavil. Miller, Galanter, and Pribram appear to regard a Plan as something "in the organism." In contrast, we regard a Plan as simply a certain kind of explanatory device. It is no more "in the organism" than the explanatory principles of geology are in the earth. In addition, we have violated the spirit, if not the letter, of Miller, Galanter, and Pribram's book by stating our Plan in the language of stimuli and responses. We defend this, if any defense is necessary, simply as a linguistic choice, and not on some ontological basis. We are not saying that behavior inherently consists of stimuli and responses; we are simply saying that the S-R language is the one in which we choose to cast our explanatory device.

A generalized Plan for serial learning.- A model embodying the aforementioned features (implicit responses and strategies) is shown in Fig. 2.

---

[1] We recognize that, in addition to the effects produced by lengthening the interval between successive responses and filling that interval with activity, errors occurring during training may be detrimental in their own right.

Figure 2.

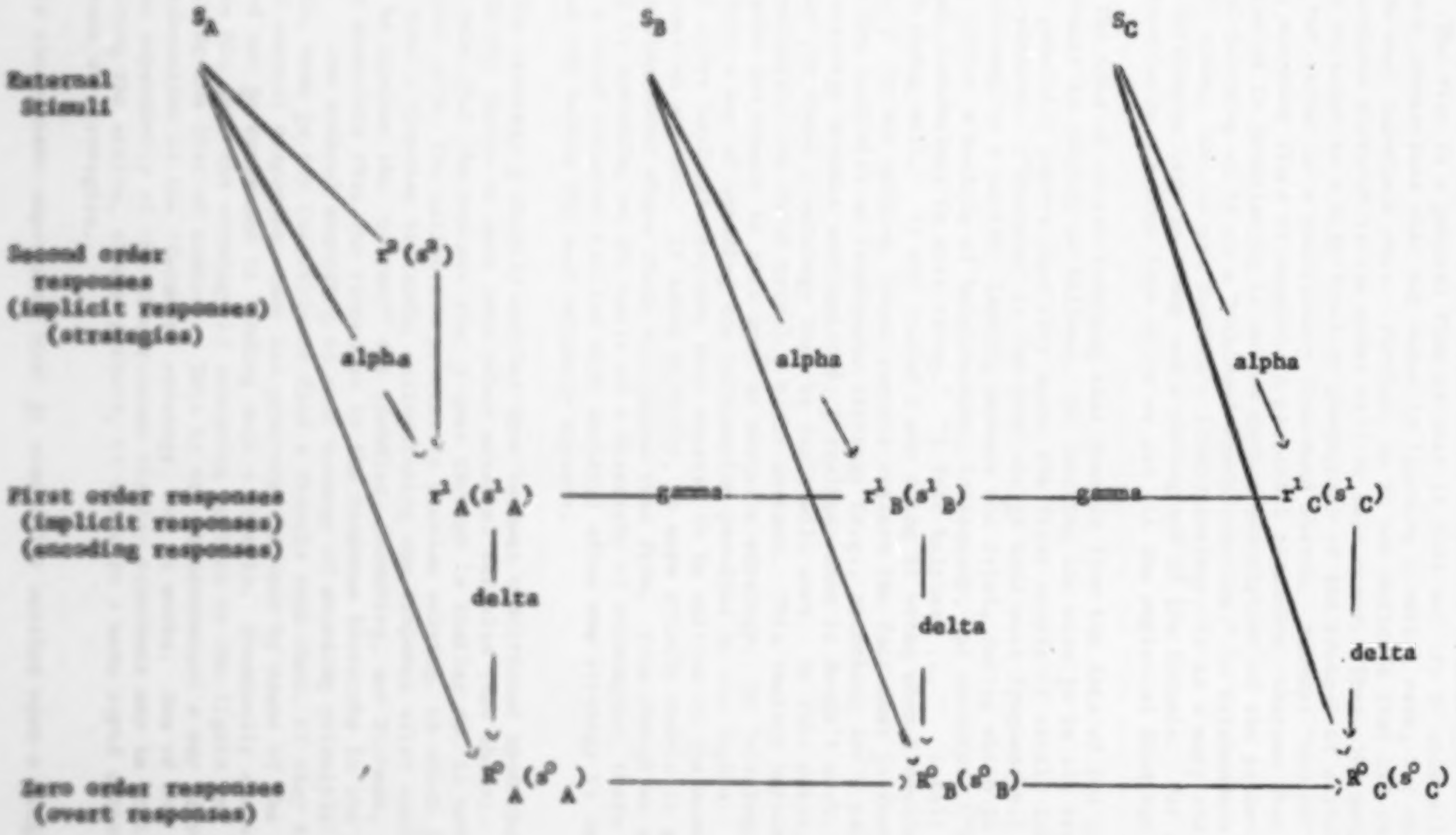

External Stimuli
Second order responses (implicit responses) (strategies)
First order responses (implicit responses) (encoding responses)
Zero order responses (overt responses)
S_A
S_B
S_C
$r^2(s^2)$
alpha
alpha
alpha
$r^1_A(s^1_A)$
gamma
$r^1_B(s^1_B)$
gamma
$r^1_C(s^1_C)$
delta
delta
delta
$R^0_A(s^0_A)$
$R^0_B(s^0_B)$
$R^0_C(s^0_C)$

Diagram of a generalized "Plan" for serial learning.

It will facilitate the explanation of the model if we assume purely for expository purposes, that the first order responses (encodings) are numerical. The Plan is a general Plan in that it does not try to show all possible S-R connections that may occur in learning a serial task, but only some of the most important ones. Further, we do not maintain that all the S-R connections pictured in the model will or must occur. Thus, the model is not to be regarded as a high-fidelity photograph of the process of serial learning, but rather as a preliminary free-hand sketch. Perhaps "sketch" is the wrong metaphor since it suggests a pictorial objective, whereas what we are interested in formulating is not so much a description of the process of serial learning as it is a "rational reconstruction," in Reichenbach's (1938) sense. Or, to use Toulmin's (1960) analogy, it is a map; and there is a difference between a map and a photograph of the terrain. Our sketch is essentially a scheme into which we can fit the empirical findings.

The idea of serial learning that emerges from the data of the present experiment is roughly as follows. In learning the maze *S*s in all treatments generally report that they spent the first couple of trials looking for a pattern. ("Pattern" is the word the *S*s used most frequently.) This was followed by a period, lasting perhaps one trial, during which *S*s typically report a feeling of helplessness, inadequacy, and anxiety. ("You can get overwhelmed by this thing." "I felt helpless ... " "I felt I wasn't doing well." "I was scared I was doing it wrong when I couldn't learn.") In our opinion, these reports reflect the fact that *S*s frequently begin the task with an inadequate strategy (e.g., searching for a pattern). This strategy becomes extinguished; *S* realizes that it doesn't work, but he has not yet found a strategy that he feels will work. At this point, (approximately the third trial) *S* feels anxious. This anxiety appears to dissipate the moment he hits upon an adequate strategy. By "strategy" we mean here a way of encoding the information provided by the lights. It is a high order implicit response that appears to be emitted in the manner of a Skinnerian operant. If asked directly, *S*s were utterly unable to give us any clue about where these strategies came from. Even though we assume that *S* is operating on the basis of a hierarchy of strategies, there seems to be a brief interval (filled with anxiety) after one strategy is extinguished and before the next strategy appears.

The strategy *S* finally settles upon becomes conditioned upon the first item in the series or upon some other external stimulus (not shown). It would seem that the process that *S* goes through is similar to, if not identical with, the extinction process in problem solving, in which *S* moves down a response hierarchy extinguishing one response after another until he reaches the "correct" one (Kendler, Greenberg, and Richman, 1952). We may speculate that the responses in the response hierarchy in the present task are ordered according to some economy of encoding principle. For example, some *S*s try initially to find a formula such that, if they know one or two correct responses, they can generate the rest by means of the formula. None of our *S*s succeeded in finding such a formula. Eventually almost all of them hit upon the strategy of assigning numbers to the lights and then memorizing the list of numbers. This is not as economical a way of encoding the information as the "formula" strategy, but it works. One of the reasons for the superiority of the Simultaneous Display treatment may be that, by exhibiting the entire, random pattern, it leads to a more rapid extinction of "formula" strategies.

In the present experiment most *S*s eventually settled upon a single strategy; that is, they encoded each stimulus in the same way: numerically. It seems quite likely that in some other serial task S might not find a single strategy adequate to encompass the entire list, and he might therefore encode, say, the first four stimuli in one way (geometrical pattern), the next three in a different way (by assigning letters), and the last five in still another way (numerically). In this case it may be hypothesized that S has a third order response, a super-ordinate implicit response that orders the strategies (first the geometrical pattern, then the letters, then the numbers). Newman and Taylor (1962) have uncovered some evidence in a paired associate experiment that supports the conjecture that an S may use several strategies in learning a task. Every strategy after the first one may be cued off by the stimulus property of the last encoding response in the set of responses governed by the preceding strategy.

From what has been said it may be noted that one role that can be assigned to the classical theory is that of explaining the sequencing of implicit responses within a strategy. A "formula" strategy yields an ordered set of responses, and hence, in this case, there is nothing for the classical theory to explain, since the sequencing is built into the strategy. However, in the case of a non-formula strategy, where the sequencing is not performed by the strategy itself, the classical theory can account for the ordering of the responses.

It is easily conceivable that serial learning might occur without the intervention of a strategy or even without implicit encoding responses. The learning could occur entirely at the level of the overt responses. This may be the way rats learn a serial task, but in the face of the evidence in the present experiment it would be hard to defend the view that this is the way human beings do it. Of course, after enough practice human Ss might be able to perform the task without mediation, but they do not seem to start off that way.

The notion that continued overt practice may lead to a bypassing of both strategies and encodings is consistent with the view of Miller, Galanter and Pribram (1960, p. 89) that the way in which an expert carries out a Plan is rather automatic. In our opinion automaticity begins when the stimulus property of each overt response begins to cue off the succeeding overt response. This points to another role that can be assigned to the classical theory. It describes, not so much the process of serial learning, as the end result of learning, after the strategies and encodings have been bypassed.

Once S has found a workable strategy he settles down to encoding the information in accordance with that strategy; e.g., saying numbers to himself. We have given the designation "alpha" to this encoding process since it seems to be the same as a process, also called "alpha," which we have reason to believe occurs in paired associate learning. (Cook and Brown, 1961).

Actually, the encoding process is not quite a simple matter of memorizing a list of numbers in which each number uniquely designates a particular light or button. For example, the eleventh button from the left is the correct one in the first row and also in the fourth row; so, encoding a button by simply designating it "11" would not enable S to discriminate between these two buttons, but other formulations would; e.g., "11 preceded by nothing" and "11 preceded by 3 and 7," or "the first 11" and "the second 11." As a further complication some Ss report that they number sometimes from the left and sometimes from the right. They tend to number from the side closest to the button. When pressed as to how they remember which side to start counting from in a particular row, Ss are at a loss to account for it.

Again, following the usage of Cook and Brown in paired associate learning, we have used "gamma" to designate the hook-up process, the process in which the stimulus property of one implicit response (or encoding) cues off the subsequent implicit response. All the evidence, both quantitative and interview data, points to the conclusion that the "rote" learning that occurs in this experiment with these Ss occurs at the level of the implicit responses rather than at the level of the overt responses. Ss in treatments in which there is no overt responding learn the maze quite easily.

Indeed, the fact that the Sequential Display Prompting treatment is significantly better than the Automatically Paced Prompting treatment points to the conclusion that at least with some Ss in some tasks -- e.g., this one -- overt practice may actually be detrimental. (The only difference between these two treatments was the presence in the APP condition of overt practice.) This finding, relatively novel in serial learning, has already been demonstrated in paired associate learning (Cook and Spitzer, 1960).

The process called "delta" in Fig. 2 -- again following Cook and Brown -- is a decoding process, a process in which the stimulus property of the encoding response evokes the overt response. In the present experiment, and probably in most laboratory serial learning tasks, this process plays little or no role because the Ss have already learned the association before they enter the experiment. They have not encoded the stimuli, and at the outset of training they probably do not even know how to do it, but they do know how to decode. If, before the experiment began, E had said to S, "Press the third button in the first row," S would have been able to do it. Both the overt response itself (button-pressing) and the connection between that response and a suitable encoding have been prelearned, and hence play no important part in this experiment.

It is quite conceivable that in another task, or with different Ss, the overt responses and the decoding processes may present a problem. For example, while no normal, literate, adult American would have any trouble in decoding the stimulus, "Draw a capital 'N'," a small child might. He might be able to pronounce "N," but not draw it. Wherever decoding is a problem, there overt practice should be beneficial. A case in point is learning to sing.

In other cases, too, overt practice might help, not because S needs practice in decoding, but because S has not been able to formulate a suitable strategy; that is, he has not been able to develop a good way of encoding the stimuli and has to fall back upon chaining the overt responses. This is conspicuously the case with serial performances that are not sustained at every point (as this one is) by external stimuli. Consider, for instance, the series of movements involved in executing a tee shot in golf. Here a series of stimuli is involved, but they are internal, response-produced stimuli that do not lend themselves well to verbal coding. The stimuli are evanescent, and we lack a vocabulary adequate to designate them.

Interpretation of the findings.- Since the outcome of the present experiment is subject to so many variables and can be looked at in so many ways, we cannot give a classically neat interpretation of the findings. However, given the particular task used in this experiment, the Plan (Fig.2) suggests some reasons why some of the experimental treatments are better than others. According to the Plan there are two important aspects to this particular task: (a) developing a strategy that works, and (b) practicing the sequencing of the encodings (gamma processes). S must, perforce, practice the encodings themselves, but we do not consider that to be important here, since the encodings themselves (numbers) are already well known to S, but the sequence in which these encodings must be arranged is novel, and not known to S prior to the experiment. We would expect that there would be an increment or decrement in performance depending upon the extent to which an experimental treatment facilitated or inhibited these two activities. Three features seem to us to be of especial importance as far as these two activities are concerned. In the first place, it seems to us that having all the stimuli on display at one time might facilitate the development of a strategy. For example, it becomes easy for S to see that the pattern is random; hence simultaneous display may tend to extinguish "formula" strategies. In the second place, giving S control over the rate at which information is fed to him may give him time to develop a strategy and permit him more opportunity to rehearse the encodings. It also allows S to decide when he wants to do the rehearsing. In the third place, forcing S to respond overtly not to get new information, but simply to practice a decoding that he already knows, means that S has less time to practice sequencing the encodings.

Looking first at the worst condition (Automatically Paced Prompting: 9.9 trials to criterion) we get some insight into why this treatment is so bad. The stimuli are displayed piecemeal, which means that a strategy that works for the first two stimuli may have to be discarded when the third one appears. But perhaps more important than this is that S has no control whatever over the rate at which information is fed to him. A new light goes on every three seconds whether he is ready for it or not. S has to be alert for each new stimulus. Moreover, S cannot simply note the new stimulus and go back to rehearsing the previous encodings (if that is what he was doing), he must respond overtly to the new light. All in all, S is kept so preoccupied with this busywork that he has little time left to devote to the things that are really important in learning this task.

The Sequential Display Prompting treatment (8.1 trials to criterion) also feeds the stimuli to S one at a time every three seconds. It keeps him alert to new stimuli, but it frees him from having to make overt responses, which is probably why it is somewhat better than the Automatically Paced Prompting condition.

The Confirmation treatment (7.3 trials), though it doles out the stimuli one at a time, does give S some control over the rate at which he receives information; he can at least slow it down by, for example, pausing when he has his finger on a correct button. Of course, S has to respond overtly in order to produce each stimulus, but these responses can be automatized if S selects the proper search strategy; e.g., always starting from the left side of each row. Indeed, if he counts to himself while pushing buttons he might hit upon the numerical strategy that much sooner.

The Sequential Display Prompting with Memory treatment (7.3) trials, while different from the Confirmation condition, produces no better results. In this treatment, though the lights come on one at a time, once they come on, they stay on. S has no control over the rate at which the lights come on. Moreover, he must keep alert for each new light since he has to respond overtly to it.

In the Simultaneous Display Prompting condition(5.9 trials) S has before him all the stimuli at the same time. He can more readily see that a formula strategy will not work, and may thus hit upon the numerical strategy sooner. To some extent, too, S has some control over the stimuli in that he can turn his attention to any part of the pattern that he wishes. Also he can practice sequencing the encodings to his heart's content.

In the Self Paced Prompting treatment (5.6 trials) S has complete control over the rate at which information is fed to him. He can go as fast or as slowly as he likes. (Ss in this treatment usually go slowly during the early trials and fast during the later ones.) There is no need for S to maintain a constant state of vigilance, and S is not distracted by the appearance of new stimuli. By controlling the information rate S has time to work out a suitable strategy and to practice sequencing the encodings. Apparently this feature compensates for the fact that all the stimuli are not on display simultaneously.

Implications for guidance.- In the light of the outcome of the present experiment it would be just as foolish to believe that any sort of guidance is better than none as to maintain that guidance is never effective. While it may be possible to devise an effective guidance technique for almost any task, two things should be done before this can be accomplished. First, as Gagné (1962) says, a task analysis should be performed. Perhaps the task is one in which the chief problem is that of decoding the stimulus properties of implicit responses that S already knows how to make. Or, perhaps the overt responses are not sustained all along the line by external stimuli, but each one must be cued off by the stimulus property of its predecessor, as in a "pure" motor skill. In cases like these an effective guidance technique might be difficult to achieve because the guidance may interfere with other cues, especially response-produced cues, upon which the sequential performance depends. For example, manual guidance in teaching handwriting to children probably introduces a drastic alteration in the cutaneous and kinesthetic cues.

However, if the analysis should reveal the task to be one in which the main problem is encoding external stimuli or developing a strategy, then it should be somewhat easier to devise an effective guidance technique, since encodings and strategies are communicable, frequently by verbal means. For example, if the path through our maze consisted of the following buttons: 6, 7, 5, 8, 4, 9, 3, 10, 2, 11, 1, 12; it would be quite easy, before S began the task or had even seen the maze, to instruct him in such a way that he would get all the responses correct on the first trial.

The second prerequisite for devising an effective guidance technique is a theory about psychological processes. Even if our task analysis has supplied us with all the facts about the task, if we have no idea of the psychological processes involved in learning the task, the formulation of a guidance procedure can only be a matter of blind trial and error. Indeed, it would appear that some people (e.g., Hanson, 1958) believe that not even an analysis can be conducted without at least a tacit theory.

Of course, no theory is wholly adequate for all purposes, and any theory may be inadequate in some respects. If a theory is inadequate in a respect that bears on the devising of a guidance technique, then that technique may not be very effective; indeed, it may be worse than no guidance at all. But if theory is not developed to the point at which it has some implications for guidance, then the formulation of a guidance technique is largely, if not entirely, a shot in the dark.

Summary

Using a 12 x 12 punchboard maze, it was found that two guidance procedures produced significantly (.01 level) faster learning than a no-guidance (confirmation) procedure, one guidance technique produced significantly (.01 level) slower learning, and two other guidance procedures yielded a learning rate that was not significantly different from that produced by the no-guidance procedure. Ss were interviewed after the experiment. Both the quantitative results and the interview data point to the explanatory inadequacy of the classical theory of serial learning, which asserts that initially each external stimulus in the list cues off an overt response, and eventually each overt response after the first one becomes conditioned upon the stimulus property of its predecessor.

In place of the classical theory the authors propose a generalized Plan, which is a mediational model of serial learning that asserts that Ss develop strategies for encoding the external stimuli, then they encoded these stimuli in accordance with the strategy, and finally, in performing the task, they decode the encodings. It is suggested that this Plan may have some explanatory power in the area of serial learning and may also prove to be of heuristic value in the development of guidance techniques for the improvement of learning.

References

Alonzo, A.S. The influence of manual guidance upon maze learning. *J. Comp. Psychol.*, 1926, 6, 143-158.

Asche, S. E., Hay, J., and Diamond, R. M. Perceptual organization in serial rote-learning. *Amer. J. Psychol.*, 1960, 73, 177-198.

Carr, H. Teaching and learning. *J. Genet. Psychol.*, 1930, 37, 189-219.

Cook, J. O., and Brown, J. E. Familiarity and novelty of stimulus and response terms in paired associate learning. *Psychol. Rep.*, 1963, 12, 535-545.

Cook, J. O., and Spitzer, M. E. Supplementary report: Prompting versus confirmation in paired associate learning. *J. exp. Psychol.*, 1960, 4, 275-276.

Ebbinghaus, H. *Uber das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Leipzig: Duncker and Humblot,1885. (Trans. as *Memory: a contribution to experimental psychology* by H. A. Ruger and C. E. Bussenius. New York: Teachers College, Columbia Univ., 1913)

Gagné, R. M. Military training and principles of learning. *Amer. Psychol.*, 1962, 17, 83-91.

Hanson, N. R. *Patterns of discovery*. Cambridge: Cambridge Univ. Press, 1958.

Hawker, J. R. The effect of prompting and confirmation in a sequential learning task. Manuscript, North Carolina State College, 1960.

Horowitz, L. M. Free recall and ordering of trigrams. *J. exp. Psychol.*, 1961, 62, 51-57.

Hull, C. L. Knowledge and purpose as habit mechanisms. *Psychol. Rev.*, 1930, 37, 511-525.

Hull, C. L. The conflicting psychologies of learning - a way out. *Psychol. Rev.*, 1935, 42, 491-516.

Hull, C. L., Hovland, C.I., Ross, R. T., Hall, M. Perkins, D. T., and Fitch F. B. *Mathematico - deductive theory of rote learning*. New Haven: Yale Univ. Press, 1940.

Kendler, H. H., Greenberg, A., and Richman, H. The influence of massed and distributed practice on the development of mental set. *J. exp. Psychol.*, 1952, 43, 21-25.

Kimble, G. A. *Hilgard and Marquis' Conditioning and Learning*. New York: Appleton-Century Crofts, 1961.

Kimble, G. A., and Wulff, J. J. "Response guidance" as a factor in the value of student participation in film instruction. In Lumsdaine, A. A. (Ed.) *Student response in programmed instruction*. Washington, D.C.:

Acad. Sci. - Natl. Resch. Coun. Publ. 943, 1961, 217-226.

Krechevsky, I. "Hypotheses" in rats. Psychol. Rev., 1932, 39, 516-532.

Lepley, W. M. A theory of serial learning and forgetting based upon conditioned reflex principles. Psychol. Rev., 1932, 39, 279-288.

Lepley, W. M. Serial reactions considered as conditioned reactions. Psychol. Monogr., 1934, 46, No. 205.

Lumsdaine, A. A., Sulzer, R. L., and Kopstein, F. F. The effect of animation cues and repetition of examples on learning from an instructional film. In Lumsdaine, A. A. (Ed.) Student response in programmed instruction, Washington, D. C.: Nat'l Acad. Sci. - Natl. Resch. Coun. Publ. 943, 1961, 241-270.

McGeoch, J. A. The psychology of human learning. 1st Ed., New York: Longmans Green, 1942.

McGeoch, J. A., and Irion, A. L. The psychology of human learning. 2nd Ed., New York: Longmans Green, 1952.

Miller, G. A., Galanter, E., Pribram, K. H. Plans and the structure of behavior. New York: Henry Holt, 1960.

Miller, N. E. Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In Koch, S. (Ed.) Psychology: A study of a science, Vol. 2, New York: McGraw-Hill, 1959.

Newman, S. E., and Taylor, L. R. Context effects in paired-associate learning as a function of element-sharing among stimulus terms. Off. Nav. Res. Tech. Rep. No. 3, Contract Nonr 486(08), 1962.

Osgood, C. E. Method and theory in experimental psychology. New York: Oxford, 1953.

Reichenbach, H. Experience and prediction. Chicago: U. of Chicago Press, 1938.

Steel, R. G. D., and Torrie, J. H. Principles and procedures of statistics. New York: McGraw-Hill, 1960.

Tolman, E. C. Cognitive maps in rats and men. Psychol. Rev., 1948, 55, 189-208.

Toulmin, S. The philosophy of science. New York: Harper, 1960.

Wulff, J. J., and Stolurow, L. M. The role of class-descriptive cues in paired-associate learning. In Lumsdaine, A. A. (Ed.) Student response in programmed instruction. Washington, D. C.: Nat'l Acad. Sci. - Natl. Resch. Coun. Publ. 943, 1961, 445-456.

SUMMARY
STUDIES IN GUIDED LEARNING

John Oliver Cook and Howard G. Miller

North Carolina State of the University of North Carolina at Raleigh
Raleigh, North Carolina

Cooperative Research Project No. 1242
July 1961 to August 1963

SUMMARY

All of the experimental studies included in this report are directly concerned in one way or another with the matter of guidance during the learning process. The first four studies are concerned with guidance and symbolic learning. The first three of these studies, A-1, A-2, and A-3, employed paired-associate learning tasks. The fourth study (A-4) is a non-experimental one which exhibits a theoretical model of paired-associate learning and proposes a method of analyzing responses. The fifth study, A-5, used a problem solving task. The sixth study, B-1, was concerned with guidance and sequential learning, and it employed a learning task. Since each study had its own specific objectives, the studies included in this report will be summarized one by one.

Experiment A-1:

Prompting and Confirmation in Paired-Associate Learning

BACKGROUND

Previous studies have shown that learning is facilitated by knowledge of results, or confirmation; that is, informing the subject about his performance after he makes the response. Other studies have shown that learning is facilitated by prompting; that is, by telling the subject which response to make before he makes it. It has been suggested that perhaps some combination of prompting and confirmation trials might prove to be superior to either one alone, and there is some evidence to support this view.

OBJECTIVES

The objective of this study was to try to find out what facilitative function, if any, is performed by confirmation trials inserted periodically in a series of prompting trials. These confirmation trials may be correctly regarded as tests.

PROCEDURE

All the subjects in this experiment were college undergraduates. The stimulus terms of the paired-associate learning task were the first ten letters of the alphabet, and the correct response was to draw a line (called a "code line") connecting three dots of a standard, non-representational, seven dot pattern. During the prompting trials, the subject saw the letter, then saw the code line, then drew the code line. A test, or in some cases, a blank interval or another prompting trial, were inserted after every three prompting trials. (A trial consisted of presenting all ten items in random order.)

For Group T_0 (N=10) there was no feedback on the test. That is, the subject was given no information about the correctness or incorrectness of his responses. Subjects in Group T_{RW}(N=10) were simply told whether their responses on the test were right or wrong. In Group T_C (N=10) after each response the subject was shown the correct response. In Group P_3 (N=5) a prompting trial instead of a test trial was inserted after every third prompting trial. For Group $(P\text{-}0)_3$ (N=5), an empty time interval equal in length to the test, was substituted for it. all subjects in this part of the experiment were given five cycles of training and test trials before the final test. In the second part of the experiment, two groups, (N=5 each) replicated the conditions of the P_3 Group and the $(P\text{-}0)_3$ Group, except that they were given only three cycles before the final test.

RESULTS

As measured by a chi-square median test, the groups given five cycles showed no significant difference in number of correct responses on the final test. Of the two groups given only three training cycles, the group with the empty time interval inserted after third prompting trial was significantly (.05 level) inferior to the group that had a prompting trial instead of an empty interval.

CONCLUSIONS

1. Given enough prompting trials, inserting or deleting tests or replacing them with additional prompting trials makes no difference (how much is enough must be determined empirically for each task).

2. As a corollary to the first conclusion, the particular combination of prompting and confirmation trials used in this experiment produced no better results than pure prompting training, even if no training trial is substituted for the confirmation trial.

3. Merely informing the subject of the correctness or incorrectness of his response on periodic test (confirmation) trials produces results that are insignificantly different from confirming his test responses by showing him the correct answer.

4. Total training time is an important variable provided it is sufficiently long.

5. If total training time is not sufficiently long, the important variable is number of trials rather than total training time.

BIBLIOGRAPHY -

There are eight references listed in the final report.

Experiment A-2

The Effect of Test Schedules Upon Learning Under Conditions of Guidance

BACKGROUND

There exists quite widespread disagreement about the optimum mixture of guided learning trials and test trials in a training program. Probably the dominant view is that the optimum schedule is one in which guided learning trials (Prompting) predominate early in learning, and test trials with feedback of the correct answer (Confirmation) predominate later in learning. However, theoretical arguments on this point are not conclusive, and empirical data are scarce.

OBJECTIVES

The objective of this study was to compare several different mixtures of guided learning trials (Prompting) and test trials with feedback of the correct answer (Confirmation).

PROCEDURE

This experiment used the same paired-associate learning materials and the same procedure that was used in Experiment A-1. The subjects were 100 college undergraduates, assigned without bias to the twenty cells of a 5 x 4 factorial design. Five different mixtures of training trials (Prompting trials early in learning, confirmation trials early in learning, regularly spaced prompting trials, pure prompting trials, and a sequence in which the subjects were allowed to choose either a prompting or a confirmation trial prior to each practice trial) were compared over four levels of training (3, 6, 9, or 12 training trials). In a follow-up experiment (N=42) additional subjects were added to each of the five foregoing treatments and, in addition, a pure confirmation group was run.

RESULTS

For both the original experiment and the follow-up study, the results, as analyzed by a non-parametric analysis of variance, showed that level of training was the only statistically significant (.01 level) variable. (For any level of training, total training time is completely confounded with number of trials.) Subjects who were given the opportunity to select their own schedule trial by trial tended to produce mixtures that most closely resembled the schedule of the prompting early group, in which prompting trials predominate early in learning, and test trials with feedback predominate later in learning. The mean scores of these two groups were the highest, though they did not differ significantly from the mean scores of the other groups.

CONCLUSIONS

1. Within the limits of the mixtures of training and test trials used in this experiment, the amount learned is not a function of the mixture.

2. The amount learned is not significantly increased or decreased by allowing the subject to select on a trial by trial basis his own mixture of training and test trials.

BIBLIOGRAPHY

There are 21 references listed in the Bibliography.

Experiment A-3

The Interacting Effects of Guidance and Anxiety Upon Learning

BACKGROUND

Previous research has tended to support the hypothesis that high-anxious subjects perform well when compared with low-anxious subjects in tasks like eyelid conditioning where there is little or no response competition, but tend to perform rather poorly in tasks like paired-associate learning where there is a great deal of response competition. One possibility is that this phenomenon is attributable, at least in part, not so much to the nature of the paired-associate task, as to the nature of the training method. Perhaps high-anxious subjects trained under a prompting procedure rather than under a confirmation procedure would not yield an inferior performance.

OBJECTIVES

1. To see whether the difference in performance between high-anxious subjects trained under a prompting procedure and high-anxious subjects trained under a confirmation procedure is greater than the difference in performance between two groups of low-anxious subjects, one group trained under prompting and the other under confirmation.

2. To determine the effect of instructionally-induced stress upon the performance of high-anxious subjects being trained under prompting and under confirmation procedures.

3. To determine the drive level as measured by number of GSR non-specifics in high-anxious subjects during training under each of the two procedures, and during testing.

PROCEDURE

Twenty male college undergraduates with scores higher than 22 on the Taylor Manifest Anxiety Scale were assigned in equal numbers to the prompting group and to the confirmation group. Twenty low-anxious subjects (scores under eight) were assigned in equal numbers to the same two training procedures. The paired-associate materials and procedure were the same as those employed in Experiment A-1. Each subject was run through five cycles, each cycle consisting of three training trials and a test.

In the second part of the study, ten high-anxious subjects were assigned to a prompting training group and ten to a confirmation group. The experimenter put both groups under stress throughout training by harshly criticizing their performance after every test. Subjects in the second part of the experiment were run through only four cycles of training. Throughout training and testing each subject's GSR was automatically monitored continuously.

RESULTS

1. In the first part of the experiment prompting was superior to confirmation for both high-and low-anxious subjects, but the difference was not statistically significant on all five tests.

2. The difference between high-anxious subjects under prompting and high-anxious subjects under confirmation was greater than the comparable difference for low-anxious subjects, but the difference between the differences was not statistically significant.

3. High-anxious subjects under stress performed better under prompting than under confirmation, but the difference is significant only for the first test.

4. Among subjects trained under prompting there is no significant difference between low-anxious subjects, high-anxious subjects, and high-anxious subjects under stress.

5. Among subjects trained under confirmation, low-anxious subjects performed better than high-anxious subjects, but the difference is not statistically significant.

6. Under confirmation high-anxious subjects under stress performed better than low-anxious subjects, and significantly better than high-anxious who were not under stress.

7. Among high-anxious subjects under stress there was very little difference in drive level, as measured by the number of GSR non-specifics, between the prompting and confirmation groups during training.

8. Test performance and level of drive during training are positively related under prompting but not related under confirmation.

9. The relation between drive level during testing and number of correct responses is positive for the prompting group and slightly negative for the confirmation group.

10. The drive level under confirmation is not significantly different from the drive level under prompting, as measured by the number of GSR non-specifics.

CONCLUSIONS

1. There is a tendency, not always statistically significant, for prompting to yield better results than confirmation, whether the subjects are high-anxious or low-anxious.

2. There is a tendency for prompting to be better than confirmation, whether or not the subjects are under stress.

3. Confirmation does not seem to generate a higher drive level than prompting.

4. There appears to be a positive relationship between drive level and performance under prompting but not under confirmation.

BIBLIOGRAPHY

There are eleven references listed in the final report.

Study A-4

Theoretical Implications of Response Analysis in Paired-Associate Learning

This non-experimental study exhibits a multiprocess model of the learning of a single paired-associate item. According to this model the following processes are involved in the learning of a single paired-associated item: the process of encoding the stimulus term (alpha process), the process of encoding the response term (beta process), the hook-up process (gamma process), and the process of decoding the encoded response term (delta process). If a subject gets a paired-associate item right on a test it indicates that all of the processes have been successfully completed. However, if he gets an item wrong, it indicates that a breakdown occurred in one or more of these processes, but we do not know which one.

The scheme for analyzing responses that is developed in this study enables us to tell, under suitable experimental conditions which are rather easy to meet, whether in the case of an incorrect item the breakdown occurred in the beta process or the gamma process. Variations of this method of response analysis should be easy to develop for serial learning tasks, and for analyzing instructional programs.

Experiment A-5

Amount of Transfer as a Function of Type of Guidance in Solving Series-Completion Problems

BACKGROUND

Everyday academic experience tends to confirm the belief that teachers are rather suspicious of the use of guidance in problem-solving tasks. It is feared that guiding a student through the solution of a problem may make the solution of a subsequent problem more difficult or may diminish his tendency to persist in trying to solve a subsequent problem.

Research in concept formation has tended to show that subjects who have formed a particular concept can form the opposite concept more easily than can subjects who did not have the opportunity to form the original concept. This shift from a concept to its opposite is called a reversal shift. To our knowledge no investigation analogous to this sort of study has been undertaken in the area of problem solving.

OBJECTIVES

1. To test the notion that in problem solving a more general sort of guidance will facilitate transfer to a new problem to a greater degree than either very specific guidance or no guidance at all.

2. To see whether guidance on one problem diminishes the tendency to persist in attempts to solve a subsequent problem.

3. To see what effect guidance has on the length of time required to solve a reversal problem.

4. To gather data on the strategies that subjects report employing in solving problems.

PROCEDURE

Thirty male undergraduates and 30 female undergraduates were assigned to the cells of a 2 x 3 factorial design in which the three treatments were specific guidance, general guidance, and no guidance. Each subject was given four problems, one at the time. Each problem required the subject to write the seventh number in a series of numbers. The dependent variable in each case was the time to solution of the problem. Where guidance was administered it was given only on the first, or training, problem. The second problem was a transfer problem. The third problem was insoluble and was used to test the persistence of the subjects. The fourth problem was a reversal of the first problem.

RESULTS

The results were analyzed (by means of nonparametric techniques) separately for each problem.

1. On problem one specific guidance was significantly better then either general guidance or no guidance.

2. On problem two there were no significant differences.

3. On problem three there were no significant differences.

4. On problem four specific guidance was significantly better than general guidance and no different from no guidance. (This outcome is attributable to the poor performance of the female-general guidance cell).

5. A few interesting, but statistically non-significant patterns emerged from an analysis of the strategies which the subjects report using.

6. There were no differences between males and females on any problem under any of the three treatments.

CONCLUSIONS

1. Guidance can be used to reduce the total time necessary to solve a series of problems.

2. Guidance on one problem has no facilitative or detrimental effect upon transfer to a similar problem.

3. Guidance on one problem has no facilitative or detrimental effect upon the tendency to persist in trying to solve a subsequent problem.

4. General guidance on one problem may have a detrimental effect upon performance on a reversal problem.

5. Extreme caution should be exercised in extending these findings to other tasks.

BIBLIOGRAPHY

There are two references listed in the final report.

Experiment B-1

A Generalized "Plan" for Serial Learning

BACKGROUND

Several studies performed in the 1920's and some in the 1950's suggest that learning can be facilitated by using guidance techniques. In the case of serial learning, this expectation is also supportable on theoretical grounds. If serial learning involves some sort of S-R chaining, then a guidance technique which reduces the length of the temporal interval between the elements of the chain as well as the amount of activity intervening between those elements should prove facilitative. However, there are no adequate theoretical grounds for predicting which kind, or kinds, of guidance techniques are likely to prove effective.

OBJECTIVES

1. To compare several different guidance procedures with a confirmation (no guidance) procedure in learning the correct path through a punchboard maze.

2. To interview subjects to try to find out how they went about learning the task.

PROCEDURES

The task was to learn the correct path through a punchboard maze which consisted of a 12 x 12 array of pushbuttons with a jewel light above each button. The path consisted of one correct button in each row. The subjects were 60 male undergraduates assigned in equal numbers without bias to the following six treatments.

1. Self paced prompting treatment (SelPP). In this treatment the light over the correct button in the first row was turned on by the experimenter. The subject then pushed the button below that light. Upon releasing the button the light over the correct button in the first row went off and the light over the correct button in the second row went on, and so on through the maze.

2. Confirmation treatment (C). Each subject pushed buttons in the first row until he hit the correct one, whereupon the light over that button came on and stayed on until the subject released the button. Then the subject did the same thing in the second row., the third row, and so on. One trip through the maze was considered one training trial.

3. Simultaneous display prompting treatment (SimDP). The lights over all 12 of the correct buttons came on and stayed on for 30 seconds. This constituted one training trial. (36 seconds was the mean time required by a pilot group of self-paced prompting subjects for one trip through the maze.) The subject made no overt button pushing responses during the training trials.

4. Sequential display prompting treatment (SeqDP). The light over the correct button in the first row automatically came on for 3 seconds and off. Then the correct light in the second row came on automatically for 3 seconds and off, and so on through the maze. The time required for one training trial was thus 36 seconds. During training the subject made no overt button pushing responses; he simply observed the lights.

5. Sequential display prompting with memory treatment (SeqDPM). This treatment was the same as the sequential display prompting treatment, except in two respects: (a) once a light came on it stayed on

until the last (12th) light went off, and (b) subjects responded overtly by pushing buttons.

6. Automatically paced prompting treatment (APP). This treatment was just like the sequential display prompting treatment except that during the 3 seconds that each light was on, the subject practiced overtly by pushing the button under that light.

All subjects were given alternate training trials and test trials without feedback until they had reached a criterion of the first of two consecutive perfect runs. After reaching criterion each subject was interviewed in an effort to find out how he went about learning the maze.

RESULTS

1. Quantitative data. In terms of number of trials to criterion, the means of the six treatments fall into three groups: (a) self-paced prompting (5.6), and simultaneous display prompting (5.9); (b) sequential display prompting with memory (7.3), confirmation (7.3), and sequential display prompting (8.1); (c) automatically paced prompting (9.9). By a parametric least significant differences test there were differences significant at the .01 level between the worst treatment in Group (a) and the best treatment in Group (b), and between the worst treatment in Group (b) and the one treatment in Group (c). There were no significant differences within these groups.

2. Interview data. Of 53 interpretable interviews, 49 subjects used a numerical strategy (assigning numbers to the buttons in each row), three subjects used a mixed numerical-spatial strategy, and one subject used a "geometric" strategy.

3. Theoretical results. The development of a theory of serial learning was an unexpected result.

CONCLUSIONS

1. Guidance as compared with no guidance is effective to a statistically significant degree in learning a punchboard maze.

2. Subjects do not appear to learn this task in a rote fashion; instead they seem to approach it by searching for a learning strategy, trying it out, and then either rejecting it or adopting it.

3. The learning behavior of the subjects is regulated by the strategy they adopt.

4. The classical (overt response chaining) theory of serial learning is inadequate to explain the data.

5. The data lend themselves to interpretation in terms of interconnections among the elements of a hierarchy of implicit responses.

PUBLICATIONS

Cook, J. O. Theoretical implications of response analysis in paired associate learning. Paper read as part of a symposium, "Theory and method in verbal learning," at the annual meeting of the Southeastern Psychological Assn., Miami Beach, Fla., 25-27 April, 1963.

Cook, J. O. 'Superstition' in the Skinnerian. *Amer. Psychol.*, 1963, 18, 516-518.

Cook, J. O., Miller, H. G., Grier, J. B., and Stamon, J. W., A generalized 'Plan' for serial learning. This paper has been accepted for publication by the *Journal of Educational Psychology*.